

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 04:56:59 ; Search time 2791.11 seconds
(without alignments)
10714.982 Million cell updates/sec

Title: US-09-367-052-3

Perfect score: 690

Sequence: 1 ctgcacctgtcagtgctga.....gctgaacccatctctat 690

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_cm.*

21: em_cr.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rod.*

36: em_hg_man.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hgto_hum.*

40: em_hgto_mus.*

41: em_hgto_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	690	100.0	1180	10	MMLCR12	Z80112 Mus musculus
2	690	100.0	1817	10	BC031665	BC031665 Mus muscu
3	690	100.0	1877	10	DB7747	DB7747 Mus muscu
4	688.4	99.8	1422	10	MMLCR13	Z80113 Mus muscu
5	685.2	99.3	1223	10	MMU59760	U99760 Mus muscu
6	685.2	99.3	3366	10	MMU65580	U65580 Mus muscu
7	682	98.8	1575	10	MMLESTRPT	X99582 M.musculu
8	682	98.8	1809	10	AB000803	AB000803 Mouse mRN
9	682	98.8	3770	10	MMLESTRGN	X99581 M.musculu
10	682	98.8	175370	2	AC145861	AC145861 Fan trogl
11	682	98.8	185215	2	AC147556	AC147556 Mus muscu
12	578.6	83.9	1050	10	AF452185	AF452185 Rattus no
13	578.6	83.9	190149	2	AC122097	AC122097 Rattus no
14	578.6	83.9	240931	2	AC107097	AC107097 Rattus no
15	575.4	83.4	1050	10	ENU90610	U90610 Rattus norv
16	523.6	75.9	1059	9	AF031089	AF031089 Papio ham
17	520.4	75.4	1044	9	AF172212	AF172212 Pygathrix
18	518.8	75.2	1003	9	AF178084	AF178084 Saguinus
19	518.8	75.2	1044	9	AF172210	AF172210 Pygathrix
20	518.8	75.2	1044	9	AF172211	AF172211 Rhinopith
21	518.8	75.2	1044	9	AF172213	AF172213 Pygathrix
22	518.8	75.2	1044	9	AF172214	AF172214 Nasalis l
23	518.8	75.2	1044	9	AF172218	AF172218 Presbytis
24	518.8	75.2	1044	9	AF172227	AF172227 Hylobates
25	518.8	75.2	1044	9	AF172228	AF172228 Hylobates
26	518.8	75.2	1044	9	AF172231	AF172231 Pongo pyg
27	518.8	75.2	1059	9	AF452613	AF452613 Saimiri s
28	517.2	75.0	901	9	AV185894	AV185894 Macaca fa
29	517.2	75.0	1003	9	AF229128	AF229128 Saimiri b
30	517.2	75.0	1004	9	AF178082	AF178082 Cercopith
31	517.2	75.0	1044	9	AF172219	AF172219 Presbytis
32	517.2	75.0	1044	9	AF172220	AF172220 Presbytis
33	517.2	75.0	1044	9	AF172221	AF172221 Mandrillu
34	517.2	75.0	1044	9	AF172223	AF172223 Macaca as
35	517.2	75.0	1044	9	AF172224	AF172224 Macaca ar
36	517.2	75.0	1044	9	AF172225	AF172225 Macaca th
37	517.2	75.0	1044	9	AF172226	AF172226 Macaca ne
38	517.2	75.0	1044	9	AF172229	AF172229 Hylobates
39	517.2	75.0	1044	9	AF172230	AF172230 Hylobates
40	517.2	75.0	1059	6	AX256185	AX256185 Sequence
41	517.2	75.0	1059	6	AX482330	AX482330 Sequence
42	517.2	75.0	1059	6	BD226095	BD226095 Therapeut
43	517.2	75.0	1059	9	AF025375	AF025375 Homo sapi
44	517.2	75.0	1059	9	AY424129	AY424129 Homo sapi
45	517.2	75.0	1059	9	BT006660	BT006660 Homo sapi

ALIGNMENTS

RESULT 1	MMLCR12	Mus musculus lcr-1 gene.	1180 bp	mRNA	linear	ROD 18-NOV-1997
LOCUS	Z80112					
DEFINITION	Z80112					
ACCESSION	Z80112.1	GI:2632100				
VERSION	Z80112.1	GI:2632100				
KEYWORDS	CXC chemokine receptor 4; CXCR-4; lcr-1 gene.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	1 (bases 1 to 1180)					
AUTHORS	Schubel, A., Burgstahler, R. and Lipp, M.					
TITLE	The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN) :					
	High expression in thymus and lymphoid tissues					

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

Unpublished
2 (bases 1 to 1180)
Lipp,M.
Direct Submission
Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
D-13122, GERMANY
On Nov 21, 1997 this sequence version replaced gi:1542888.
Location/Qualifiers
1..1180
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129SV"
/db_xref="taxon:10090"
/tissue_type="thymus"
/dev_stage="adult"
1..1180
/gene="lcr-1"
1..1074
/gene="lcr-1"
/function="receptor for chemokine SDF-1"
/standard_name="CXC chemokine receptor 4"
/codon_start=1
/product="CXCR-4"
/protein_id="CAB02202.1"
/db_xref="GI:1542889"
/db_xref="GOA:P70658"
/db_xref="SWISS-PROT:P70658"
/translation="MEPIGIYSDNYSEVSGSDYDSNKEPCFRDENVHENRIFLPTI
YFIPLTGVGLVGLVWYQKLRSDTKYRLHLSVADLLFVTLPLFAVADAMADW
YFGKFLCAVHIYITVNLVSSVLLAFISLDRIYLAIVHATNSQRPRLKLAKEANTVGV
WPAALLTIPDFIFADVSQSDISQGDYICDRPLDLSLMMVVFQFQHMVGLILPFI
VILSCYIIISKLHSGKQKALKTLTILAFACWLPYVIGISDLSILLGVK
QGDFFSIVHKWISITLALAFHCLNPLIYFLGAKFKSSAQHALNSMRGSLKIL
SKGKRGHSSVSTESSESSPHSS"
<1..15
/gene="lcr-1"
/number=1
16..>1180
/gene="lcr-1"
/number=2

exon

exon

ORIGIN

Query Match 100.0%; Score 690; DB 10; Length 1180;
Best Local Similarity 100.0%; Fred. No. 7.2e-163; Indels 0; Gaps 0;
Matches 690; Conservative 0; Mismatches 0

QY 1 CTGCACCTGTACAGCGGTTCTCATCTCGGCTTCATCAGCGTGGACCGGTACCTCGCC 180
DB 232 CTGCACCTGTACAGCGGTTCTCATCTCGGCTTCATCAGCGTGGACCGGTACCTCGCC 411
QY 61 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCAATCATCTACACT 120
DB 292 GCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCAATCATCTACACT 351
QY 121 GTCAACCTCTACAGCGGTTCTCATCTCGGCTTCATCAGCGTGGACCGGTACCTCGCC 180
DB 352 GTCAACCTCTACAGCGGTTCTCATCTCGGCTTCATCAGCGTGGACCGGTACCTCGCC 411
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCGCAAGAACTCTGGCTGAAAGCGATCTAT 240
DB 412 ATTGTCCAGCCACCAACAGTCAAGGCGCAAGAACTCTGGCTGAAAGCGATCTAT 471
QY 241 GTGGGGCTGTGGATCCAGCGCTTCCTCTGACTATACCTGACTTCTATCTTTCCGACGTC 300
DB 472 GTGGGGCTGTGGATCCAGCGCTTCCTCTGACTATACCTGACTTCTATCTTTCCGACGTC 531
QY 301 AGCCAGGGGACATCATAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
DB 532 AGCCAGGGGACATCATAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 591
QY 361 AGCCTGTGATGGTGGTGTTCATTCACGATATATATGTTGGTGGTCTATCTGCGCCGC 420

Db 592 AGCCTGTGGATGGTGGTGTTCATTCACGATATATATGTTGGTCTCATCTCTGCCGCG 651
QY 421 ATGCTCATCT 480
Db 652 ATGCTCATCT 711
QY 481 CAGAAGCGCAGGCCCTCTCAAGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 712 CAGAAGCGCAGGCCCTCTCAAGACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 771
QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 772 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 831
QY 601 GGATGTGACTTCGAGACATGTGCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 832 GGATGTGACTTCGAGACATGTGCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 891
QY 561 TTCCACTGTTTGCCTGAACCCCATCTCTAT 690
Db 892 TTCCACTGTTTGCCTGAACCCCATCTCTAT 921

RESULT 2
BC031665
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BC031665 1817 bp mRNA linear ROD 07-OCT-2003
Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone
MGC:36266 IMAGE:3592479), complete cds.
BC031665
BC031665.1 GI:21618781
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1817)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Sheerz,T.E., Brownstein,M.J., Utsid,T.B., Toshikiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Aramon,R.D., Mullany,S.J., Bosak,S.A., McWhan,F.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S.D., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1817)
Strausberg,R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Baylor College of Medicine Human Genome

Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgobcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 61 Row: d Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: similarity but not
identity to protein.

FEATURES
source
1. .1817
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36265 IMAGE:3592479"
/tissue_type="mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP Mam1"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1. .1817
/gene="Cxcr4"
/note="synonyms: LESTR, fusin, PB-CXR, PBSF/SDF-1, CD184"
/db_xref="LocusID:12767"
/db_xref="MGI:109563"
78. .1157
/codon_start=1
/product="chemokine (C-X-C motif) receptor 4"
/protein_id="AAH31685.1"
/db_xref="GI:21618782"
/db_xref="LocusID:12767"
/translation="MESPISVTSNYSSEVSGVDYDSNKEPCFRDENVHFNRLP
TIVFLITGVGNGLVILVMGYOKRLVMTDKVRLHSLVADLLFVITLPPAVADAMA
DWYKFLCAVHIIVVNLVSYLLIAPISLDRLAIHATNSQRPRLKAPKAVV
GWTPALLITIPFIPDRVYQSGDISQDDRYICDRLYPSDLNVMVFPQHIWGLILP
GIVLSCYCIILKLSHSGQRKALKTVILILAFACWLPYVIGVISIDSLFLLGV
ILKSGPDESIVHKWISTEALAFPHCLNPLIYAFILGAFKFAQHALNSMRGSLK
246. .1004
/note="7tm_1; Region: 7 transmembrane receptor (rhodopsin
family)"
/db_xref="CDD:pfam00001"

ORIGIN

Query Match 100.0%; Score 690; DB 10; Length 1817;
Best Local Similarity 100.0%; Pred. No. 7.3e-163;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCACCTGTCAGTGGTCACTCTCTTGTCTATCACTCCCTCTGGGAGTTGAT 60
Db 315 CTGCACCTGTCAGTGGTCACTCTCTTGTCTATCACTCCCTCTGGGAGTTGAT 374
Qy 61 GCCATGGCTGACTGTTACTTGGGAATTTTGTGAAGCTGTCCATATCATCTACCT 120
Db 375 GCCATGGCTGACTGTTACTTGGGAATTTTGTGAAGCTGTCCATATCATCTACCT 434
Qy 121 GTCAACCTTACAGCAGGTTCTCATCTGCGCTTTCATAGCTGGACCGGTACCTGCC 180
Db 435 GTCAACCTTACAGCAGGTTCTCATCTGCGCTTTCATAGCTGGACCGGTACCTGCC 494
Qy 181 ATTGTCCAGCCCAACAGTCAAGGCCAAGAACTGCTGCTGAAGAGGAGTCTAT 240
Db 495 ATTGTCCAGCCCAACAGTCAAGGCCAAGAACTGCTGCTGAAGAGGAGTCTAT 554
Qy 241 GTGGGCGTGTGATCCAGCCCTCTCTGACTATACCTGACTTCATCTTTGCGGAGTC 300
Db 555 GTGGGCGTGTGATCCAGCCCTCTCTGACTATACCTGACTTCATCTTTGCGGAGTC 614

Qy 301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTATCATCTGTGACCGCTTTACCCCGAT 360
Db 615 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTATCATCTGTGACCGCTTTACCCCGAT 674
Qy 361 AGCCTGTGATGGTGGTGGTTCATTAATTCAGCATATAATGGTGGTCTCATCTGCGCGGC 420
Db 675 AGCCTGTGATGGTGGTGGTTCATTAATTCAGCATATAATGGTGGTCTCATCTGCGCGGC 734
Qy 421 ATCTGTCATCTCTCTGTGTACTGTCATCATCTCTAAGCTGTACACTTCCAAAGGGCCAC 480
Db 735 ATCTGTCATCTCTCTGTGTACTGTCATCATCTCTAAGCTGTACACTTCCAAAGGGCCAC 794
Qy 481 CAGAGCGCAGGCGCTCAAGACGACAGTCACTCTCATCTCTAGCTTTCTTTGCGCTGTGG 540
Db 795 CAGAGCGCAGGCGCTCAAGACGACAGTCACTCTCATCTCTAGCTTTCTTTGCGCTGTGG 854
Qy 541 CTGCGCATATTATGTGGGATCAGCATCGACTCTCTCATCTCTTTGCGGAGTCAATCAAGAA 600
Db 855 CTGCGCATATTATGTGGGATCAGCATCGACTCTCTCATCTCTTTGCGGAGTCAATCAAGAA 914
Qy 601 GGATGTGACTTCGAGAGCATTGTGCACAAAGTGGATCTCCATCAGAGGCGCTTCGCTTC 660
Db 915 GGATGTGACTTCGAGAGCATTGTGCACAAAGTGGATCTCCATCAGAGGCGCTTCGCTTC 974
Qy 661 TTCCACTGTTCCTGAACCCCATCTCTAT 690
Db 975 TTCCACTGTTCCTGAACCCCATCTCTAT 1004
RESULT 3
D87747
LOCUS D87747 1877 bp mRNA linear ROD 07-FEB-1999
DEFINITION Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION D87747
VERSION D87747.1 GI:1772445
KEYWORDS murine CXCR-4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Nagasaki, T., Nakajima, T., Tachibana, K., Iizasa, H., Bleul, C.C.,
Yoshie, O., Matsushima, K., Yoshida, N., Springer, T.A. and
Kishimoto, T.
Molecular cloning and characterization of a murine pre-B-cell
growth-stimulating factor/stromal cell-derived factor-1 receptor, a
murine homolog of the human immunodeficiency virus 1 entry
coreceptor fusin
Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
97121456
8962122
2 (bases 1 to 1877)
Nagasaki, T.
Molecular cloning of murine PBSF/SDF-1 receptor
Unpublished
3 (bases 1 to 1877)
Nagasaki, T.
Direct Submission
Submitted (05-SEP-1996) Takashi Nagasaki, Research Institute, Osaka
Medical Center, department of immunology; Murodoko 840, Izumi,
Osaka 590-02, Japan (E-mail: immunol@osk.threewebnet.or.jp.
Tel:0725-56-1220, Fax:0725-57-3021)
Location/Qualifiers
1. .1877
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="DW34"
/cell_type="pre-B cell"
/tissue_type="bone marrow"
1. .1877
/gene="murine CXCR-4"

CDS

120..1199
/gene="murine CXCR-4"
/function="PDSF/SDF-1 receptor"
/codon_start=1
/evidence=experimental
/product="murine CXCR-4"
/protein_id="BAA13451.1"
/db_xref="GI:1772446"
/translation="MEPIVSIVTSDNYSEEYSGDYSNKEPCFRDENVHNRIFLP
TYIFIFLTCIVGVNGLIVMVGQKLSMTDKYRLHLSVADLLFVILPFWADVAMA
DNVFGFLCKAVHIIYVNLVSGVLILAFISLDRLVLAIVHATNSORPKRLAEKAVV
GWIPALLATIPDFADYQSDISOGDDRYICDRLYPSLWMMVVFQFQHMVGLILP
GIVILSCYIILKSHSGHQRKALKTIVILAFACWLPYVYVIGISIDSPILGV
IKQGCDFESIVHKWISLIPALAPFHCCLNPIIFYAFGLAKFKSSAHLNMSRGSLLK
ILSKRGKRGHSSVSTESSESPHSS"

ORIGIN

Query Match 100.0%; Score 690; DB 10; Length 1877;
Best Local Similarity 100.0%; Pred. No. 7,3e-163;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGACCTGTACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 60
DB 357 CTGACCTGTACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 416
QY 61 GCCATGGCTGACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 120
DB 417 GCCATGGCTGACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 476
QY 121 GTCAACCTCTACAGCAGCGTTCTATCTGCGCTTCATCAGCTGACCGGTACCTCGCC 180
DB 477 GTCAACCTCTACAGCAGCGTTCTATCTGCGCTTCATCAGCTGACCGGTACCTCGCC 536
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAGGACGCTCTAT 240
DB 537 ATTGTCCAGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAGGACGCTCTAT 596
QY 241 GTGGGGCTGTGATCCAGCGCTCTCTGACATATACCTGACTTCATCTTTGCCGACGTC 300
DB 597 GTGGGGCTGTGATCCAGCGCTCTCTGACATATACCTGACTTCATCTTTGCCGACGTC 656
QY 301 AGCCAGGGGACATCATGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 360
DB 657 AGCCAGGGGACATCATGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 716
QY 361 AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTGTCTATCTGCGCCGCG 420
DB 717 AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTGTCTATCTGCGCCGCG 776
QY 421 ATCGTCATCTCTCTGTACTGATCATCATCTCTAAGCTGTCACTCCAGGCGCCAC 480
DB 777 ATCGTCATCTCTCTGTACTGATCATCATCTCTAAGCTGTCACTCCAGGCGCCAC 836
QY 481 CAGAAGCGGAGCGCTCAAGACGACATCTCTCATCTAGCTTTCTTTGCTGCTGG 540
DB 837 CAGAAGCGGAGCGCTCAAGACGACATCTCTCATCTAGCTTTCTTTGCTGCTGG 896
QY 541 CTGCCATATTATGTGGGGATCAGCATCGATCTCTCATCTTTTGGAGTCTATCAAGCAA 600
DB 897 CTGCCATATTATGTGGGGATCAGCATCGATCTCTCATCTTTTGGAGTCTATCAAGCAA 956
QY 601 GGAATGTGACTTCAGAGCATTTGCAAGTGTATCTCCATCAGAGCGCTTCGCTTC 560
DB 957 GGAATGTGACTTCAGAGCATTTGCAAGTGTATCTCCATCAGAGCGCTTCGCTTC 616
QY 661 TTCACCTGTGTGCTGAAACCCCTCTCTAT 690
DB 1017 TTCACCTGTGTGCTGAAACCCCTCTCTAT 1046

MMLCR13 1422 bp DNA linear ROD 31-OCT-1996

RESULT 4

MMLCR13
LOCUS

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus lcr-1 gene, exon 2.
Z80113
Z80113.1 GI:1655638
CXCR chemokine receptor 4; CXCR-4; lcr-1 gene.
Mus musculus (house mouse)
Mus musculus

REFERENCE

1 (bases 1 to 1422)
Schubel, A., Burgstahler, R. and Lipp, M.
The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
High expression in thymus and lymphoid tissues
Unpublished
2 (bases 1 to 1422)
Lipp, M.

AUTHORS
TITLE
JOURNAL

Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
Molecular Medicine, MDC, Robert-Roessle-Strasse 10, BERLIN-BUCH,
D-13122, GERMANY
Location/Qualifiers
1..1422

FEATURES

source
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129SV"
/db_xref="taxon:10090"
/clone_lib="lambda fix (TM) II"
/dev_stage="adult"
/number=1
260..1422
/gene="lcr-1"
260..>1422
/gene="lcr-1"
/product="CXCR-4"
/standard_name="CXCR chemokine receptor 4"
/number=2
/function="receptor for chemokine SDF-1"

ORIGIN

Query Match 99.8%; Score 688.4; DB 10; Length 1422;
Best Local Similarity 99.9%; Pred. No. 1.8e-162;
Matches 689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGACCTGTACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 60
DB 476 CTGACCTGTACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 535
QY 61 GCCATGGCTGACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 120
DB 536 GCCATGGCTGACGTGCTGACCTCCCTCTTTGTGATCACACTCCCTCTCGGCGAGTTGAT 595
QY 121 GTCAACCTCTACAGCAGCGTTCTATCTGCGCTTCATCAGCTGACCGGTACCTCGCC 180
DB 596 GTCAACCTCTACAGCAGCGTTCTATCTGCGCTTCATCAGCTGACCGGTACCTCGCC 655
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAGGACGCTCTAT 240
DB 656 ATTGTCCAGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAGGACGCTCTAT 715
QY 241 GTGGGGCTGTGATCCAGCGCTCTCTGACATATACCTGACTTCATCTTTGCCGACGTC 300
DB 716 GTGGGGCTGTGATCCAGCGCTCTCTGACATATACCTGACTTCATCTTTGCCGACGTC 775
QY 301 AGCCAGGGGACATCATGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 360
DB 776 AGCCAGGGGACATCATGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 835
QY 361 AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTGTCTATCTGCGCCGCG 420
DB 836 AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTGTCTATCTGCGCCGCG 895
QY 421 ATCGTCATCTCTCTGTACTGATCATCATCTCTAAGCTGTCACTCCAGGCGCCAC 480

Db	896	ATCGTCACTCTCTCTCTGTTACTGTCATCATCATCTCTTAAGCTGTGCACACTCCAAAGGGCCAC	955			
Qy	481	CAGAAGCGCAAGGCCCTCAAGACACACAGTCATCTCTCATCTCTAGCTTTCTTTTGCTGCTGG	540			
Db	956	CAGAAGCGCAAGGCCCTCAAGACACACAGTCATCTCTCATCTCTAGCTTTCTTTTGCTGCTGG	1015			
Qy	541	CTGCCATATTATGTGGGATCAGCATCGACTCCTTTCATCTCTTTTGGAGTCATCAAGCAA	600			
Db	1016	CTGCCATATTATGTGGGATCAGCATCGACTCCTTTCATCTCTTTTGGAGTCATCAAGCAA	1075			
Qy	601	GGATGTGACTTTCGAGAGCAATGTGCACAAAGTGGATCTCCATCAAGAGGCGCTCGCCTTC	660			
Db	1076	GGATGTGACTTTCGAGAGCAATGTGCACAAAGTGGATCTCCATCAAGAGGCGCTCGCCTTC	1135			
Qy	661	TTCCACTGTTGCCCTGAACCCCACTCTAT	690			
Db	1136	TTCCACTGTTGCCCTGAACCCCACTCTAT	1165			
RESULT 5						
MMU59760						
LOCUS	MMU59760	1223 bp	linear			
DEFINITION	Mus musculus fusin homolog mRNA, complete cds.					
ACCESSION	U59760					
VERSION	U59760.1	GI:1527134				
KEYWORDS						
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
	1 (bases 1 to 1223)					
	Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.					
	Cloning of the mouse fusin gene, homologue to a human HIV-1					
JOURNAL	co-factor					
	J. Immunol. 157 (12), 5455-5460 (1996)					
	97113334					
	8955194					
	2 (bases 1 to 1223)					
REFERENCE	Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.					
AUTHORS	Direct Submission					
TITLE	Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200					
JOURNAL	Longwood Avenue, Boston, MA 02115, USA					
FEATURES	Location/Qualifiers					
	1..1223					
	/organism="Mus musculus"					
	/mol_type="mRNA"					
	/strain="129/Sv"					
CDS	/db_xref="taxon:10090"					
	30..1109					
	/note="seven transmembrane spanning receptor;					
	chemokine-receptor like"					
	/codon_start=1					
ORIGIN	/product="fusin homolog"					
	/protein_id="AA007725.1"					
	/db_xref="GI:1527135"					
	/translations="MPSIVSVITYSDNYSZEEVSGSDYDNKSPCFRDENVHNRIFLP					
	TYIFILFTIGVINGNGLVLMVGQKRLSMTDKRLHLSVADLLFVITLPFFWADVAMA					
Query Match	DWYFGFLCKAVHIYIADYVSVLLAFISLDYLAIVHATNSQRPKLLAEKAVY					
	GWVPALLATVPIITFADYVSOQDISQDDRYICDRLYPDLSLWVWQFQHIVGLIPL					
	GIVLSCYCIILSKLSGHQKRLKALKTVLILFAFCWLPPYVVGISIDSFILGLV					
	IKQCFDSIVHKWISITELAFPHCLNPLIYAFILGAKFKSAQHALLNSMRGSSLK					
	ILSKRGHGHSSVSTSESSSFSS"					
Query Match 99.3%; Score 685.2; DB 10; Length 1223;						
Best Local Similarity 99.6%; Pred. No. 1.2e-161;						
Matches 687; Conservative 0; Mismatches 3; Indels 0; Gaps 0;						
Qy	1	CTGCACCTGTCACTGGCTGACCTCTCTTTGTTCATCACACTCCCTCTGGGAGTTGAT	60			
Db	267	CTGCACCTGTCACTGGCTGACCTCTCTTTGTTCATCACACTCCCTCTGGGAGTTGAT	326			
Qy	61	GCCATGGCTGACCTGGTACTTTTGGGAATTTTGTGTAAGGCTGTCCATATCATCTACACT	120			

```
gene      1..3366
          /gene="CXCR-4"
          join(1..21,2308..3366)
CDS       /gene="CXCR-4"
          /note="SDF-1 chemokine receptor"
          /codon_start=1
          /product="fusin"
          /protein_id="AAC52953.1"
          /db_xref="GI:3645915"
          /translation="MEPISVISIYTSNDYSEVSGDYDSNKEPCFRDENHFNRIPLP
          TYFIITGTVGNGLVILVMGYOKLRSMYDKYRLHLSVADLLFVITLFPVADVAMA
          DWYFKGFKCAVHIITVNLVSYSLILAFSLDRYLAIVHATNSQRPKLLAEKAVYV
          GVMIPALLITPDPIDFADVSQDIDISQDDRYICDRLYPDSLMVMVFOFHIMVGLILP
          GIVILSCVCIILSKLSHKGHOKRKALKTTVILILAFPAFCWLPVYVGISIDSFILGV
          IKQGDPSIVHKWISITEALAPFHCCNPLILAFILGAKFKYSSAQHALNSMRGSSLK
          ILSKGRGSHSVSTESSESFHSS"
exon      1..21
          /gene="CXCR-4"
          /number=1
exon      1..15
          /gene="CXCR-4"
          /note="alternate exon 1"
intron    16..>2307
          /gene="CXCR-4"
          /note="alternate intron 1"
intron    22..2307
          /gene="CXCR-4"
          /number=1
exon      2308..3366
          /gene="CXCR-4"
          /number=2

ORIGIN
Query Match      99.3%; Score 685.2; DB 10; Length 3366;
Best Local Similarity 99.6%; Pred. No. 1.2e-161;
Matches 687; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGCACCTGTGAGTGGTGAACCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 60
DB 2524 CTGCACCTGTGAGTGGTGAACCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 2583
QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACT 120
DB 2584 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACT 2643
QY 121 GTCACCTCTACAGAGCGTTTCTCATCTGCGCTTTCATCAGCTGACCGGTACCTCGCC 180
DB 2644 GTCACCTCTACAGAGCGTTTCTCATCTGCGCTTTCATCAGCTGACCGGTACCTCGCC 2703
QY 181 ATTGTCACGCGCCACACAGTCAAGGCCAGGAACTGCTGGCTGAAGGCGAGTCTAT 240
DB 2704 ATTGTCACGCGCCACACAGTCAAGGCCAGGAACTGCTGGCTGAAGGCGAGTCTAT 2763
QY 241 GTGGCGTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTTGGCGAGTC 300
DB 2764 GTGGCGTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTGCGGAGTC 2823
QY 301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
DB 2824 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 2883
QY 361 AGCCTGTGATGTGGTGTCTTTCAATTCAGCATATATATGGTGGGTCTCATCTTCCCGCGC 420
DB 2884 AGCCTGTGATGTGGTGTCTTTCAATTCAGCATATATATGGTGGGTCTCATCTTCCCGCGC 2943
QY 421 ATGCTCATCTCTCTCTGTACTGATCATCATCTTAAGTGTGCACACTCCAGGGGCGAC 480
DB 2944 ATGCTCATCTCTCTCTGTACTGATCATCATCTTAAGTGTGCACACTCCAGGGGCGAC 3003
QY 481 CAGAAGCGCAAGGGCCCTCAAGCACAGTCACTCTCTCTAGCTTCTTTTCTGCTGCTGG 540
DB 3004 CAGAAGCGCAAGGGCCCTCAAGCACAGTCACTCTCTCTAGCTTCTTTTCTGCTGCTGG 3063
```

```
QY 541 CTGCCATATTATGTGGGATCAGCATCAGTCCCTTCATCTCTTTGGGAGTATCAAGCAA 600
DB 3064 CTGCCATATTATGTGGGATCAGCATCAGTCCCTTCATCTCTTTGGGAGTATCAAGCAA 3123
QY 601 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTTCATCACAGAGCCCTCGCCCTTC 660
DB 3124 GGATGTGACTTCGAGAGCATTGTGCACAAGTGGATCTTCATCACAGAGCCCTCGCCCTTC 3183
QY 661 TTCCACTGTTGCTGAACCCCATCTCTAT 690
DB 3184 TTCCACTGTTGCTGAACCCCATCTCTAT 3213

RESULT 7
MMLESTRPT 1575 bp mRNA linear ROD 11-NOV-1996
LOCUS M.musculus mRNA for leukocyte-derived seven transmembrane domain
DEFINITION
ACCESSION X99582.1 GI:1666648
VERSION X99582
KEYWORDS lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Moepps,B., Frodl,R., Kessler,H. and Gierschik,P.
cDNA cloning and genomic organization of a leukocyte-derived seven
transmembrane domain receptor (LESTR) from mouse: a murine
homologue of the human HIV-1 entry cofactor fusin
Unpublished
REFERENCE 2 (bases 1 to 1575)
AUTHORS Moepps,B.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,
Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK Revised by author 11-NOV-96
COMMENT On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES
Location/Qualifiers
1..1575
/organism="Mus musculus"
/mol_type="mRNA"
/strain="B6CBA"
/db_xref="taxon:10090"
/haplotype="diploid"
/sex="female"
/tissue_type="thymus"
/dev_stage="6-8 weeks old"
1..1575
/gene="lestr"
30..1159
/gene="lestr"
/codon_start=1
/product="leukocyte-derived seven transmembrane domain
receptor"
/protein_id="CAA67894.1"
/db_xref="GI:1666649"
/db_xref="GGA:P70658"
/db_xref="SWISS-PROT:P70658"
/translation="MEPISVISIYTSNDYSEVSGDYDSNKEPCFRDENHFNRIPLP
TYFIITGTVGNGLVILVMGYOKLRSMYDKYRLHLSVADLLFVITLFPVADVAMA
DWYFKGFKCAVHIITVNLVSYSLILAFSLDRYLAIVHATNSQRPKLLAEKAVYV
GVMIPALLITPDPIDFADVSQDIDISQDDRYICDRLYPDSLMVMVFOFHIMVGLVLP
GIVILSCVCIILSKLSHKGHOKRKALKTTVILILAFPAFCWLPVYVGISIDSFILGV
IKQGDPSIVHKWISITEALAPFHCCNPLILAFILGAKFKYSSAQHALNSMRGSSLK
ILSKGRGSHSVSTESSESFHSS"
polya_signal
1498..1503
/gene="lestr"

ORIGIN
Query Match      98.8%; Score 682; DB 10; Length 1575;
Best Local Similarity 99.3%; Pred. No. 7.5e-161;
Matches 685; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

QY 1 CTGCACTGTAGTGGCTGACCTCTCTTTGTATCATCACTCCCTCTTCTGGGCACTTGTAT 60
 Db |||||
 327 CTGCACTGTAGTGGCTGACCTCTCTTTGTATCATCACTCCCTCTTCTGGGCACTTGTAT 386
 QY 61 GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAAGGTGTCATATCATCTACT 120
 Db |||||
 387 GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAAGGTGTCATATCATCTACT 446
 QY 121 GTCACCTCTACAGCAGCGTTCTCATCTCGCCCTTCTATCGCTGGACCGGTACTCGCC 180
 Db |||||
 447 GTCACCTCTACAGCAGCGTTCTCATCTCGCCCTTCTATCGCTGGACCGGTACTCGCT 506
 QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAACCTGTGGCTGAAAGGCGAGTCTAT 240
 Db |||||
 507 ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAACCTGTGGCTGAAAGGCGAGTCTAT 566
 QY 241 GTGGCGCTCTGGATCCAGCCCTCTCTCTGACTACTCTGACTTCACTTCTTTCGCGAGTC 300
 Db |||||
 567 GTGGCGCTCTGGATCCAGCCCTCTCTCTGACTACTCTGACTTCACTTCTTTCGCGAGTC 626
 QY 301 AGCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGAT 360
 Db |||||
 627 AGCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGAT 686
 QY 361 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGTCTCATCTGCCCGGC 420
 Db |||||
 687 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGTCTCATCTGCCCGGC 746
 QY 421 ATCGTATCTCTCTGTTACTGATCATCATCTCTTAAGTGTGCACACTCCAAGGGCCAC 480
 Db |||||
 747 ATCGTATCTCTCTGTTACTGATCATCATCTCTTAAGTGTGCACACTCCAAGGGCCAC 806
 QY 481 CAGAAGCGCAGGCGCTCAGACACAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 Db |||||
 807 CAGAAGCGCAGGCGCTCAGACACAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 866
 QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
 Db |||||
 867 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 926
 QY 601 GGATGTGACTTCGAGAGCATGTGCACAGTGTGATCTCCATCAAGAGGCCCTCGCCTTC 660
 Db |||||
 927 GGATGTGACTTCGAGAGCATGTGCACAGTGTGATCTCCATCAAGAGGCCCTCGCCTTC 986
 QY 661 TTCCACTGTTGCCCTGAACCCCATCTCTAT 690
 Db |||||
 987 TTCCACTGTTGCCCTGAACCCCATCTCTAT 1016

RESULT 8
 AB000803
 LOCUS Mouse mRNA for murine CXCR-4, complete cds. 1809 bp mRNA linear ROD 05-FEB-1999
 DEFINITION
 ACCESSION AB000803
 VERSION AB000803.1 GI:1816445
 KEYWORDS murine CXCR-4; fusin.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1809)
 Suzuki, G., Nakata, Y., Uzawa, A., Shirasawa, T., Saito, T. and Mita, K.
 Molecular cloning of murine CXCR-4, a murine homologue of HIV entry
 co-receptor, that is expressed on leukocyte and other organs
 Unpublished
 2 (bases 1 to 1809)
 Suzuki, G.
 Direct Submission
 Submitted (29-JAN-1997) Gen Suzuki, National Institute of
 Radiological Sciences, Division of Radiation Health; 4-9-1,
 Anagawa, Inage-ku, Chiba, Chiba 263, Japan
 (E-mail:gsuzuki@numx1.nirs.go.jp, Tel:043-251-2111,

Fax: 043-284-1736) Location/Qualifiers
 1..1809
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL-6"
 /db_xref="taxon:10090"
 /tissue_type="thymus"
 /dev_stage="E17"
 72..1151
 /note="fusin"
 /codon_start=1
 /product="murine CXCR-4"
 /protein_id="BAA19187.1"
 /db_xref="GI:1816446"
 /translation="MEPISVSTYTDNYSSEVSGDYSDNSKEPCFRDENVHNRIFLP
 TYIFLFTGIVGNGLVILVMSYOKKLSMTDKYRLHLSVADLLFVITLFWAVDAM
 DMYFKCKAVHIIYTNLYSSMLLAFISLDRLVLAIVHATNSORPKLAEKAVYV
 GWYIPALLTTPDFIADVSQGDISQGDRIICDRLYEDSLWMVYFQGHINVLVLP
 GIVLSCIIISKLSHQKRAKALITVILIAFFACWLPFYVGLSIDSFILGV
 INQGDPSIVHKWISITEALAFHCCCLNPILYAFLGAKFKSSAQAHLNSMRGSLK
 ILSKGRGSHSSVSESSSFHSS"

ORIGIN
 Query Match 98.8%; Score 682; DB 10; Length 1809;
 Best Local Similarity 99.3%; Pred. No. 7,6e-161;
 Matches 685; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CTGCACTGTAGTGGCTGACCTCTCTTTGTATCATCACTCCCTCTTCTGGGCACTTGTAT 60
 Db 309 CTGCACTGTAGTGGCTGACCTCTCTTTGTATCATCACTCCCTCTTCTGGGCACTTGTAT 368
 QY 61 GCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAAGGTGTCATATCATCTACT 120
 Db 369 GCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAAGGTGTCATATCATCTACT 428
 QY 121 GTCAACCTCTACAGCAGCGTTCTCATCTCGCCCTTCTATCGCTGGACCGGTACTCGCC 180
 Db 429 GTCAACCTCTACAGCAGCGTTCTCATCTCGCCCTTCTATCGCTGGACCGGTACTCGCT 488
 QY 181 ATGTGTCACGCCACCAACAGTCAAGGCCAAGGAACCTGTGGCTGAAAGGCGAGTCTAT 240
 Db 489 ATGTGTCACGCCACCAACAGTCAAGGCCAAGGAACCTGTGGCTGAAAGGCGAGTCTAT 548
 QY 241 GTGGCGCTCTGGATCCAGCCCTCTCTCTGACTACTACTCTGACTTCACTTTCGCGAGTC 300
 Db 549 GTGGCGCTCTGGATCCAGCCCTCTCTCTGACTACTACTCTGACTTCACTTTCGCGAGTC 608
 QY 301 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGAT 360
 Db 609 AGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGAT 668
 QY 361 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGTCTCATCTGCCCGGC 420
 Db 669 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGTCTCATCTGCCCGGC 728
 QY 421 ATCGTATCTCTCTGTTACTGATCATCATCTCTTAAGTGTGCACACTCCAAGGGCCAC 480
 Db 729 ATCGTATCTCTCTGTTACTGATCATCATCTCTTAAGTGTGCACACTCCAAGGGCCAC 788
 QY 481 CAGAAGGCCAAGGCCCTCAAGACAGATCATCTCTAGTCTTCTTTTGGCTGCTGG 540
 Db 789 CAGAAGGCCAAGGCCCTCAAGACAGATCATCTCTAGTCTTCTTTTGGCTGCTGG 848
 QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
 Db 849 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 908
 QY 601 GGATGTGACTTCGAGAGCATGTGCACAGTGTGATCTCCATCAAGAGGCCCTCGCCTTC 660
 Db 909 GGATGTGACTTCGAGAGCATGTGCACAGTGTGATCTCCATCAAGAGGCCCTCGCCTTC 968
 QY 661 TTCCACTGTTGCCCTGAACCCCATCTCTAT 690

Db 969 TTCCACTGTGCTGAACCCATCTCTAT 998

RESULT 9

MMLESTRGN

LOCUS

DEFINITION

M.musculus gene encoding leukocyte-derived seven transmembrane domain receptor, strain B6.

ACCESSION

X99581.1 GI:1924959

VERSION

lestr gene; leukocyte-derived seven transmembrane domain receptor.

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Moepps, B., Frodl, R., Kessler, H. and Gierschik, P.

CDNA cloning and genomic organization of a leukocyte-derived seven

transmembrane domain receptor (LESTR) from mouse: a murine

homologue of the human HIV-1 entry cofactor fusin

Unpublished

REFERENCE

2 Moepps, B.

Direct Submission

Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm,

Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG

Revised by [3]

3 (bases 1 to 3770)

REFERENCE

Moepps, B.

Direct Submission

Submitted (02-APR-1997) B. Moepps, Universitaet Ulm,

Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG

On Apr 3, 1997 this sequence version replaced gi:1666646.

COMMENT

FEATURES

Location/Qualifiers

1..3770

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="B6"

/db_xref="taxon:10090"

/haplotypes="diploid"

/tissue type="thymus"

/dev stage="6-8 weeks old"

<1..110

/number=1

join(90..110,2306..3364)

/gene="lestr"

join(90..110,2306..3364)

/gene="lestr"

/codon start=1

/product="leukocyte-derived seven transmembrane domain

receptor"

/protein_id="CAA67893.1"

/db_xref="GI:1666647"

/db_xref="GOA:P70658"

/db_xref="SWISS-PROT:P70658"

/translation="MEPTSVSYTSDNYSEEVSGDYDSNKEPCFRDENVHNRIFLP

TIYFIITGIVGGLVILVNGYOKLRSMTDKYELHLSVADLLFVITLPPWADAVA

DWYFGKFLCKAVHIIYTNLVSSVLLAFISLDRIYLAIVATNSQRPKLLAEKAVIV

GWIPALLTLTPDFADVSQDLSQGDRIYICDRLYPDSLWMVYFQFHIMVGLVLP

GIVLSCYIIISKLSHSGHKQKALKITVILAFACWLPYVIGSIDSFILGV

IKQGDFFSIYVHWISITAEALAFHCLNPLVYAFGLAKFKFSQAQHALNSMRGSLK

ILSKGRGHSVSSTESSESFHSS"

105..2305

/gene="lestr"

/note="splice variant/intron"

111..2305

/gene="lestr"

/number=1

2306..3370

/number=2

3693..3698

misc_feature

intron

exon

polyA_signal

ORIGIN

Query Match

Best Local Similarity

Matches 685; Conservative

98.8%; Score 682; DB 10; Length 3770;

99.3%; Pred. NO. 7.8e-161;

0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTGACCTGTGAGTGCTGACCTCTCTTTGTTCATCACACTCCCTCTCTGGGAGTGTAT 60

Db 2522 CTGACCTGTGAGTGCTGACCTCTCTTTGTTCATCACACTCCCTCTCTGGGAGTGTAT 2581

QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACT 120

Db 2582 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACT 2641

QY 121 GTCAACCTCTACAGCAGCGTTCTCATCCGGCTTCATCAGCTGAGCGGTACCTCGGC 180

Db 2642 GTCAACCTCTACAGCAGCGTTCTCATCCGGCTTCATCAGCTGAGCGGTACCTCGGT 2701

QY 181 ATTGTCCAGCCCAACAGTCAAAGGCCAAGAACTCTGGCTGAAAGGAGTCTAT 240

Db 2702 ATTGTCCAGCCCAACAGTCAAAGGCCAAGAACTCTGGCTGAAAGGAGTCTAT 2761

QY 241 GTGGGCTGTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCATCTTCCGACGTC 300

Db 2762 GTGGGCTGTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCATCTTCCGACGTC 2821

QY 301 AGCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360

Db 2822 AGCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 2881

QY 361 AGCTGTGATGGTGTTCATTCAGCAGTAAATGGTGGTCTCATCTGCCCCGC 420

Db 2882 AGCTGTGATGGTGTTCATTCAGCAGTAAATGGTGGTCTCATCTGCCCCGC 2941

QY 421 ATCGTCATCTCTCTGTGTACTGTCATCATCTCTTAAGCTGTACACTTCCAAAGGCCAC 480

Db 2942 ATCGTCATCTCTCTGTGTACTGTCATCATCTCTTAAGCTGTACACTTCCAAAGGCCAC 3001

QY 481 CAGAAGCGCAAGCCCTCAAGACGACGTCATCTCACTAGCTTCTTGGCTGTCTGG 540

Db 3002 CAGAAGCGCAAGCCCTCAAGACGACGTCATCTCACTAGCTTCTTGGCTGTCTGG 3061

QY 541 CTGCCATATTATGTGGGATCAGATCGACTCTCTTTCATCTTGGAGTCTATCAAGCAA 600

Db 3062 CTGCCATATTATGTGGGATCAGATCGACTCTCTTTCATCTTGGAGTCTATCAAGCAA 3121

QY 601 GGATGTGACTTCGAGAGCATTTGTGACAAAGTGGATCTCCATCACAGAGCCCTCGCTTC 660

Db 3122 GGATGTGACTTCGAGAGCATTTGTGACAAAGTGGATCTCCATCACAGAGCCCTCGCTTC 3181

QY 661 TTCCACTGTGCTGACCCCATCTCTAT 690

Db 3182 TTCCACTGTGCTGACCCCATCTCTAT 3211

RESULT 10

AC145861

LOCUS

DEFINITION

Pan troglodytes chromosome UNK clone RP43-21D9, *** SEQUENCING IN

PROGRESS ***, 39 unordered pieces.

AC145861

AC145861.1 GI:33386791

ACCESSION

VERSION

KEYWORDS

HTG; HTGS PHASE1.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE

1 (bases 1 to 175370)

AUTHORS

Wilson, R.K.

TITLE

The sequence of Pan troglodytes clone

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 175370)

AUTHORS

Wilson, R.K.

TITLE

Direct Submission

Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: C_PT021D09

* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```

1      5345: contig of 5345 bp in length
*      5346: gap of unknown length
*      5446: contig of 11750 bp in length
*      17196: gap of unknown length
*      17296: contig of 15955 bp in length
*      33251: gap of unknown length
*      33351: contig of 12296 bp in length
*      45647: gap of unknown length
*      45747: contig of 16246 bp in length
*      61993: gap of unknown length
*      62093: contig of 1331 bp in length
*      63424: gap of unknown length
*      63524: contig of 1143 bp in length
*      64667: gap of unknown length
*      64667: gap of unknown length
*      64767: contig of 1228 bp in length
*      65995: gap of unknown length
*      66095: contig of 1336 bp in length
*      67431: gap of unknown length
*      67531: contig of 1452 bp in length
*      68983: gap of unknown length
*      69083: contig of 1809 bp in length
*      70892: gap of unknown length
*      70992: contig of 1181 bp in length
*      72173: gap of unknown length
*      72273: contig of 1342 bp in length
*      73615: gap of unknown length
*      73715: contig of 1471 bp in length
*      75186: gap of unknown length
*      75286: contig of 1569 bp in length
*      76855: gap of unknown length
*      76955: contig of 1459 bp in length
*      78414: gap of unknown length
*      78514: contig of 1451 bp in length
*      79965: gap of unknown length
*      80065: contig of 1681 bp in length
*      81746: gap of unknown length
*      81846: contig of 1572 bp in length
*      83418: gap of unknown length
*      83518: contig of 1393 bp in length
*      84911: gap of unknown length
*      85011: contig of 1778 bp in length
*      86789: gap of unknown length
*      86889: contig of 2004 bp in length
*      88893: gap of unknown length
*      88993: contig of 1678 bp in length
*      90671: gap of unknown length
*      90771: contig of 2437 bp in length
*      93208: gap of unknown length
*      93308: contig of 3351 bp in length
*      96659: gap of unknown length
*      96759: contig of 3818 bp in length
*      100577: gap of unknown length
*      100677: contig of 4312 bp in length
*      104989: gap of unknown length

```

```

*      105089: contig of 3805 bp in length
*      108994: gap of unknown length
*      112632: contig of 3639 bp in length
*      112732: gap of unknown length
*      116328: contig of 3596 bp in length
*      116329: gap of unknown length
*      116428: contig of 4552 bp in length
*      120580: gap of unknown length
*      120581: contig of 5741 bp in length
*      121081: gap of unknown length
*      126822: contig of 4410 bp in length
*      126922: gap of unknown length
*      131332: contig of 7274 bp in length
*      131432: gap of unknown length
*      131432: contig of 7274 bp in length
*      138706: gap of unknown length
*      145169: contig of 6364 bp in length
*      145170: gap of unknown length
*      145270: contig of 5427 bp in length
*      150697: gap of unknown length
*      150797: contig of 9192 bp in length
*      159989: gap of unknown length
*      160089: contig of 7156 bp in length
*      167245: gap of unknown length
*      167345: contig of 8026 bp in length.

FEATURES
      source
      1..175370
      /organism="Pan troglodytes"
      /mol_type="genomic DNA"
      /db_xref="taxon:9598"
      /chromosome="UNK"
      /clone="RP43-21D9"
      1..5345
      /note="assembly_name:Contig100"
      5446..17195
      /note="assembly_name:Contig101"
      17296..33250
      /note="assembly_name:Contig102"
      33351..45646
      /note="assembly_name:Contig103"
      45747..61992
      /note="assembly_name:Contig104"
      62093..63423
      /note="assembly_name:Contig16"
      63524..64666
      /note="assembly_name:Contig3"
      64767..65994
      /note="assembly_name:Contig45"
      66095..67430
      /note="assembly_name:Contig48"
      67531..68982
      /note="assembly_name:Contig5"
      69083..70891
      /note="assembly_name:Contig56"
      70992..72172
      /note="assembly_name:Contig60"
      72273..73614
      /note="assembly_name:Contig67"
      73715..75185
      /note="assembly_name:Contig68"
      75286..76854
      /note="assembly_name:Contig69"
      76955..78413
      /note="assembly_name:Contig73"
      78514..79964
      /note="assembly_name:Contig75"
      80065..81745
      /note="assembly_name:Contig77"
      81846..83417
      /note="assembly_name:Contig79"
      83518..84910
      /note="assembly_name:Contig80"
      85011..86788
      /note="assembly_name:Contig81"
      86889..88892

```


2738 5059: contig of 2322 bp in length
 5060 5159: gap of unknown length
 6204 6203: contig of 1044 bp in length
 6304 6303: gap of unknown length
 11814 11814: contig of 5511 bp in length
 11914 11914: gap of unknown length
 18453 18453: contig of 6539 bp in length
 18554 18554: gap of unknown length
 33685 33685: contig of 15131 bp in length
 33785 33785: gap of unknown length
 57714 57714: contig of 23929 bp in length
 57814 57814: gap of unknown length
 84552 84552: contig of 26738 bp in length
 84652 84652: gap of unknown length
 122286 122286: contig of 37634 bp in length
 122386 122386: gap of unknown length
 185215 185215: contig of 62830 bp in length.

FEATURES
 source
 1..185215
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP23-480N14"
 1..1119
 /note="assembly_name:Contig1951"
 1220..2637
 /note="assembly_name:Contig103"
 2738..5059
 /note="assembly_name:Contig122"
 5160..6203
 /note="assembly_name:Contig124"
 6304..11814
 /note="assembly_name:Contig126"
 11915..18453
 /note="assembly_name:Contig127"
 18554..33684
 /note="assembly_name:Contig128"
 33785..57713
 /note="assembly_name:Contig129"
 57814..84551
 /note="assembly_name:Contig130"
 84652..122285
 /note="assembly_name:Contig131"
 122386..185215
 /note="assembly_name:Contig132"

ORIGIN
 Query Match 98.8%; Score 682; DB 2; Length 185215;
 Best Local Similarity 99.3%; Pred. No. 9e-161;
 Matches 685; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 CTGACCTGTACGTGGTACCTCTTTTGTTCATCACACTCCCTTCTGGGAGTTGAT 60
 91537 CTGACCTGTACGTGGTACCTCTTTTGTTCATCACACTCCCTTCTGGGAGTTGAT 91596
 61 GCCATGCTGACTGGTACTTTGGGAATTTTGTGAAGCTGTCCATATCATCTACACT 120
 91597 GCCATGCTGACTGGTACTTTGGGAATTTTGTGAAGCTGTCCATATCATCTACACT 91656
 121 GTCAACCTGTACAGCAGCTTCTATCTGGCCCTTCATAGCTGACCGGACTTCGCC 180
 91657 GTCAACCTGTACAGCAGCTTCTATCTGGCCCTTCATAGCTGACCGGACTTCGCT 91716
 181 ATTGTCCACGCCACCAACAGCTCAAGGCCAAGAACTGTGCTCAAAAGCAGTCTAT 240
 91717 ATTGTCCACGCCACCAACAGCTCAAGGCCAAGAACTGTGCTCAAAAGCAGTCTAT 91776
 241 GTGGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCTATCTTTCCGACGTC 300
 91777 GTGGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCTATCTTTCCGACGTC 91836
 301 AGCCAGGGGACATCAGTCAGGGGGATGACAGTACATCTGTGACCGCCTTTACCCCGAT 360

91837 AGCCAGGGGACATCAGTCAGGGGGATGACAGTACATCTGTGACCGCCTTTACCCCGAT 91896
 361 AGCCTGTGATGGTGGTGTTCATTAATTCAGCATATAAATGGTGGTCTCATCTGCCCGC 420
 91897 AGCCTGTGATGGTGGTGTTCATTAATTCAGCATATAAATGGTGGTCTCATCTGCCCGC 91956
 421 ATCGTCACTCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTCACTCCAAAGGCCAC 480
 91957 ATCGTCACTCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTCACTCCAAAGGCCAC 92016
 481 CAGAAGCGCAAGGCCCTCAAGACGACGTCATCTCTAGCTGTCTTTCCTGCTGTGG 540
 92017 CAGAAGCGCAAGGCCCTCAAGACGACGTCATCTCTAGCTGTCTTTCCTGCTGTGG 92076
 541 CTGCCATATTATGTTGGGATCAGCATCGACTCTCTCATCTTTTGGGAGTCAACAAGCAA 600
 92077 CTGCCATATTATGTTGGGATCAGCATCGACTCTCTCATCTTTTGGGAGTCAACAAGCAA 92136
 601 GGATGTGACTTCGAGAGCATTTGTCACAAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
 92137 GGATGTGACTTCGAGAGCATTTGTCACAAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 92196
 661 TTCCACTGTTCGCTGAACCCCATCTCTAT 690
 92197 TTCCACTGTTCGCTGAACCCCATCTCTAT 92226

RESULT 12
 AF452185
 LOCUS
 DEFINITION
 Rattus norvegicus strain Holtzman chemokine receptor CXCR4 (Cxcr4)
 gene, complete cds.
 ACCESSION
 AF452185
 VERSION
 AF452185.1 GI:17902280
 KEYWORDS
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 1050)
 Simen, A.A. and Miller, R.J.
 TITLE
 Chemokine regulation of neuronal signaling and gp120 neurotoxicity
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 1050)
 Simen, A.A. and Miller, R.J.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (25-NOV-2001) Dept. Psychiatry, Yale University, New Haven, CT 06508, USA
 FEATURES
 Location/Qualifiers
 1..1050
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Holtzman"
 /db_xref="taxon:10116"
 /sex="female"
 /tissue_type="whole brain"
 1..1050
 /gene="Cxcr4"
 1..1050
 /gene="Cxcr4"
 /codon_start=1
 /product="chemokine receptor CXCR4"
 /protein_id="AA147855.1"
 /db_xref="GI:17902281"
 /translation="MEIYSDNYSSEVSGDYDSNKEPCPRDENENRIFLPTIYFI
 IFLTGVNGLVLMGYQKLRMTDKYRLHLVADLLFVITLPFWAVDAMDWYFG
 KFLCAVHIITVNLVSLVILAFISLDRLVLAIVHATNSORPKLLAEKAVYVGMIP
 ALLVTIPIIDVPSQSDGRIYCDRLVPSLWVWVFOFHIMVGLIIPGIVILSCYCI
 ILSKLSHSGHKQKALKTKTIVILAFACWLPVYVIGSIDSFILLEVIKQGEPEFV
 VHWISITEALAFPHCCINFLIYAFLGAKFKSSAQHAINSMRGSLSKILSKRGKRGH
 SSVSTESSESSFSHSS"

variation 429..430
/gene="Ckcr4"
/note="compared to Wistar strain sequence presented in
GenBank Accession Number U9610"
/replaces="cg"

ORIGIN

Query Match 83.9%; Score 578.6; DB 10; Length 1050;
Best Local Similarity 91.4%; Pred. No. 8e-135;
Matches 630; Conservative 0; Mismatches 44; Indels 15; Gaps 1;

QY 1 CTGCACTGTGAGTGGTGAAGTCTCTCTTTGTCATCACTCCCTCTTGGGCGATGAT 60
DB 223 CTGCACTGTGAGTGGTGAAGTCTCTCTTTGTCATCACTCCCTCTTGGGCGATGAT 282
QY 61 GCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACT 120
DB 283 GCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACT 342
QY 121 GTCAACCTCTAGCAGAGGCTTCTATCTCTGCTTCTATCAGCTGGAGCGGTACTCGCC 180
DB 343 GTCAACCTCTAGCAGAGGCTTCTATCTCTGCTTCTATCAGCTGGAGCGGTACTCGCC 402
QY 181 ATTGTCCAGCCCAACAGTCAAGGCAAGGAACTGCTGGCTGAAAGCGATCTAT 240
DB 403 ATTGTCCAGCCCAACAGTCAAGGCAAGGAACTGCTGGCTGAAAGCGATCTAT 462
QY 241 GTGGGGCTGTGATCCAGCCCTCTCTCTGATATATACCTGATCTCTTTGCCAGCTC 300
DB 463 GTGGGGCTGTGATCCAGCCCTCTCTCTGATATATACCTGATCTCTTTGCCAGCTC 522
QY 301 AGCCAGGGGGATCAGTCAGGGGATGACAGGTATCTGTGACCGCTTTACCCGAT 360
DB 523 AGCCAGGGGGATCAGTCAGGGGATGACAGGTATCTGTGACCGCTTTACCCGAT 567
QY 361 AGCTGTGTGATGGTGGTGTTCATTCAGCAGTATATATGGTGGTGTCTCATCTCGCCGCG 420
DB 568 AGCTGTGTGATGGTGGTGTTCATTCAGCAGTATATATGGTGGTGTCTCATCTCGCCGCG 627
QY 421 ATCGTCATCTCTCTCTGTACTGATCATCTCTATAGCTGTGATCTCAAGGCGAC 480
DB 628 ATCGTCATCTCTCTGTACTGATCATCTCTATAGCTGTGATCTCAAGGCGAC 687
QY 481 CAGAGCGCAAGCGCTTCAAGCAGCAGTCACTCTCATCTCTAGCTTTCTTGTCTGTCTGG 540
DB 688 CAGAGCGCAAGCGCTTCAAGCAGCAGTCACTCTCATCTCTAGCTTTCTTGTCTGTCTGG 747
QY 541 CTGCCATATATGTGGGATGATGATCTCTCTCTCTCTCTCTTTTGGGAGTATCAAGCAA 600
DB 748 CTACCGTATTAGCTGGGATGATGATCTCTCTCTCTCTCTTTTGGGAGTATCAAGCAA 807
QY 601 GGATGTGCTTCGAGAGCATTTGTGACAGTGGATCTCCATCAGAGCGCTTCCCTTC 660
DB 808 GGATGTGCTTCGAGAGCATTTGTGACAGTGGATCTCCATCAGAGCGCTTCCCTTC 867
QY 661 TTCCACTGTGCTGAACCCATCTCTTA 699
DB 868 TTCCACTGTGCTGAACCCATCTCTTA 896

RESULT 13
AC122097
LOCUS
DEFINITION Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC122097 190149 bp DNA linear HTG 15-NOV-2002
VERSION AC122097.4 GI:25008263
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 190149)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devila, W.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiadis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaatane, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, I., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jarkson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Miloslavjevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Paoletti, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pastorak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL

Unpublished
2 (bases 1 to 190149)
Worley, K.C.

REFERENCE
AUTHORS

Direct Submission
Submitted (22-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 190149)
Rat Genome Sequencing Consortium.

TITLE
JOURNAL

Unpublished
2 (bases 1 to 190149)
Worley, K.C.

REFERENCE
AUTHORS

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 15, 2002 this sequence version replaced gi:23269619.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence

QY	361	ACCTGTGGATGGTGGTGGTTCCTCAATTCAGCATATATATGGTGGGTCTCATCTCCCGGC	420
Db	21223	ACCTGTGGATGGTGGTGGTTCCTCAATTCAGCATATATATGGTGGGTCTCATCTCCCGGC	21282
QY	421	ATCGTCATCCCTCTCTCTGTACTGTGATCATCATCTCTAAGCTGTGCACACTCAAGGGGCAC	480
Db	21283	ATCGTCATCCCTCTCTGTACTGTGATCATCATCTCTAAGCTGTGCACACTCAAGGGGCAC	21342
QY	481	CAGAAGCGCAAGGCCCTCAAGACGACAGTCATCTCTCATCTCTAGTCTCTTTGGGAGTCATCAAGCA	540
Db	21343	CAGAAGCGCAAGGCCCTCAAGACTACGCTCATCTCTATCTCTGGCTCTCTTTGGCTGTGG	21402
QY	541	CTGCCATATTATGTGGGATCAGATCAGCTACCTCTCTCTCTCTCTTTGGGAGTCATCAAGCA	600
Db	21403	CTACCGTATTACGTGGGATCAGATCAGCTCTCTCTCTCTCTCTTTGGAGTCATCAAGCA	21462
QY	601	GGATGTGACTTCGAGGACATTTGTGCAAGGTGATCTCCATCAGAGGCCCTCGCCTTC	660
Db	21463	GGATGTGAGTTTCGAGAGCGTGTGCAAGGTGATCTCCATCAGAGGCCCTCGCCTTC	21522
QY	661	TTCCACTGTTGGCTGAACCCCATCTCTTA	689
Db	21523	TTCCACTGTTGGCTGAACCCCATCTCTTA	21551
RESULT 14			
AC107097/c			
LOCUS	AC107097	240931 bp	DNA linear HTG 13-MAY-2003
DEFINITION	Rattus norvegicus clone CH230-175010, *** SEQUENCING IN PROGRESS ***		
ACCESSION	AC107097		
VERSION	AC107097.5	GI:30580624	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 240931)		
AUTHORS	Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chaves, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyne, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Demarco, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabis, A., Gant, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyak, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuخوا, L., Louissegh, H., Lozada, R. J., Lu, X., Ma, J., Mangeshwar, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mayhoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoilemeh, O., Okwunodu, G., Olarinpinuagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plonner, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,		

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Stearns, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 240931)
Worley, K.C.

Submitted (16-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 240931)
Rat Genome Sequencing Consortium.

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23115278.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKSK
Center clone name: CH230-175010
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 216793 bases at least Q40
Consensus quality: 219254 bases at least Q30
Consensus quality: 221224 bases at least Q20
Estimated insert size: 228277; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 240931: contig of 240931 bp in length.
Location/Qualifiers
1..240931
/organism="Rattus norvegicus"

FEATURES
source

		/mol_type="genomic DNA"	
		/db_xref="taxon:10116"	
		/clone="CH230-175010"	
misc_feature	1..1249	/notes="wgs_end_extension"	
		clone_end:Sp6"	
		complement(5161..5587)	
misc_feature		/notes="clone boundary"	
		clone_end:Sp6"	
		site:EcORI	
		end sequence: BH357332"	
misc_feature		complement(232579..233220)	
		/notes="clone boundary"	
		clone_end:T7"	
		site:EcORI	
		end sequence: BH357330"	
misc_feature		235468..237558	
		/notes="wgs_end_extension"	
		clone_end:T7"	
misc_feature		239450..240931	
		/notes="wgs_end_extension"	
		clone_end:T7"	
ORIGIN			
	Query Match	83.9%; Score 578.6; DB 2; Length 240931;	
	Best Local Similarity	91.4%; Pred. No. 9.8e-135; Indels 15; Gaps 1;	
	Matches 630; Conservative 0; Mismatches 44;		
Qy	1	CTGCACCTGTACAGTGGCTGACCTCCCTTTGTGTCATCACACTCCCTCTCTGGCAGTGGAT	60
Db	237401	CTGCACCTGTGCGTGGCTGACCTCCCTTTGTGTCATCACACTCCCTCTCTGGCAGTGGAC	237342
Qy	61	GCCATGGCTACAGTGGTACCTTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACT	120
Db	237341	GCCATGGCTACAGTGGTACCTTTGGGAAATTTTGTGTAAGGCTGTGATATCATCTACACC	237282
Qy	121	GTCAACCTCTACAGCAGCGTTCTCATCTGGCCCTTCATCAGCCTGGACCGGTACCTCGCC	180
Db	237281	GTCAACCTTTACAGCAGTGTCTCATCTGGCCCTTCATCAGCCTGGACCGCTACCTTGCC	237222
Qy	181	ATTGTCCACGCCACCAACAGTCAAGCCCAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTAT	240
Db	237221	ATTGTCCACGCCACCAACAGTCAAGCCCAAGAACTGCTGCTGCTGCTGCTGCTGCTAT	237162
Qy	241	GTGGGCTGTGGATCCCGCCCTCTCTGACATATACCTGACATCTCTTTGCGGACGTC	300
Db	237161	GTGGGCTGTGGATCCCGCCCTCTCTGACATATACCTGACATCTCTTTGCGGACGTC	237102
Qy	301	AGCCAGGGGACATCATAGTCAGGGGATGACAGTATCTGTGACCGCCCTTTACCCGAT	360
Db	237101	AGCCAGGGGACATCATAGTCAGGGGATGACAGTATCTGTGACCGCCCTTTACCCGAC	237057
Qy	361	AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTCTCATCTGCGCGGC	420
Db	237056	AGCTGTGGATGTGGTGTTCATATCCAGCATATATGTTGGTCTCATCTGCGCGGC	236997
Qy	421	ATCGTATCTCTCTCTGTACTGTCATCATCTCTTAAGCTGTCACTCCAAAGGGCCAC	480
Db	236996	ATCGTATCTCTCTCTGTACTGTCATCATCTCTTAAGCTGTCACTCCAAAGGGCCAC	236937
Qy	481	CAGNAGCGGAGGCGCTCAAGACAGCAGTCATCTCTAGCTTTCTTTGCTGCTG	540
Db	236936	CAGNAGCGGAGGCGCTCAAGACAGCAGTCATCTCTAGCTTTCTTTGCTGCTG	236877
Qy	541	CTGCCATTTATGTGGGATCAGCATCGACTCCCTTCATCTCTTTGGGAGTCATCAAGCAA	600
Db	236876	CTACCGTATTACGTGGGATCAGCATCGATTCTCTCTTTGGGAGTCATCAAGCAA	236817
Qy	601	GGATGTGACTTCAGAGCATTGTGCAAGTGTGATCTTCATTCAGAGGCCCTTCGCTTC	660
Db	236816	GGATGTGACTTCAGAGCATTGTGCAAGTGTGATCTTCATTCAGAGGCCCTTCGCTTC	236757
Qy	661	TTCCACTGTTGCTGACACCCCATCTCTCTA	689

Db 236756 TTCCACGTGTCCTGAACCCCACTCTTA 236728

RESULT 15
 RN090610
 LOCUS
 DEFINITION RN090610 1050 bp mRNA linear ROD 26-MAR-1997
 Rattus norvegicus CXc chemokine receptor (CXCR4) mRNA, complete cds.

ACCESSION U90610
 VERSION U90610.1 GI:1906612
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 1050)
 AUTHORS Harrison,J.K. and Salafraanca,M.N.
 TITLE Molecular cloning of rat CXCR4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1050)
 AUTHORS Harrison,J.K. and Salafraanca,M.N.
 TITLE Direct Submission
 JOURNAL Submitted (25-FEB-1997) Pharmacology and Therapeutics, University of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA

FEATURES
 source
 1..1050
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /tissue_type="spleen"
 1..1050
 /genes="CXCR4"
 1..1050
 /genes="CXCR4"
 /notes="G-protein coupled receptor"
 /codon_start=1
 /product="CXc chemokine receptor"
 /protein_id="AAB50408.1"
 /db_xref="GI:1906613"
 /translations="MRYISDNYSEVSGDYDSNKPCFRDENENPNRIFLPIYFI
 IFUTGNGVGLVTLVWGYYQKQSMTDKYRLHLSVADLLFVITLPWADVADWADWTFG
 KLFCKAVHIIYVNLVSVLLIFSLDIYLAIVHATNSQARKLAELAEKAVYGVWIP
 ALLTLPDIIFADVSGDGRYICDRLYPDLSLMWVFQHIWVGLLPLGIVILSCVCI
 IISLGHSGHQKRAKLTVIIILAFACWLPPYVVGISIDSPILLEVTKQCGEPCSI
 VHKWISITRALAPFHCLNPLIYVFLGAKFKSSAQHALNSMRGSSLSKILSKGRGGH
 SSVSTSESSSFHSS"

ORIGIN

Query Match 83.4%; Score 575.4; DB 10; Length 1050;
 Best Local Similarity 91.1%; Pred. No. 5.1e-134;
 Matches 628; Conservative 0; Mismatches 46; Indels 15; Gaps 1;

QY 1 CTGCACCTGTCAGCTGGACCTCCCTTTGTTCATCACACTCCCTCTCTGGCAGTTGAT 60
 Db 223 CTGCACCTGTCCTGGCTGACCTCCCTCTTTGTTCATCACACTCCCTCTCTGGCAGTTGAC 282

QY 61 GCCATGGCTGACCTGGTACTTTGGGAAATTTTGTGAAGGCTGCCATATCATCTACACT 120
 Db 283 GCCATGGCTGACCTGGTACTTTGGGAAATTTTGTGAAGGCTGTCATATCATCTACACC 342

QY 121 GTCAACCTCTACAGCAGCGTTCTCATCTGGCCCTTCATAGCCTGGACCGTACTCGCC 180
 Db 343 GTCAACCTTTACAGCAGTGTCTCATCTGGCCCTTCATAGCCTGGACCGCTACTTGGC 402

QY 181 ATTGTCCACGCCACCAACAGCTCAAAGCCAAAGAAATGCTGGCTGAAAAGGCAGTCTAT 240
 Db 403 ATTGTCCACGCCACCAACAGCAGCGGAGGAGCTGCTGGCTGAAAAGGCCGCTAT 462

QY 241 GTGGCGCTTGGATCCAGCCCTCTCTGATATACCTGACTTCATCTTTGCGCAGCTC 300

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 03:46:58 ; Search time 295.894 Seconds
(without alignments)
9906.453 Million cell updates/sec

Title: US-09-367-052-3

Perfect score: 690

Sequence: 1 ctgcacctgtcagtggtga.....gcctgaacccatctctat 690

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	690	100.0	1758	2	Aaz27611 Mouse CXCR4
2	690	100.0	1877	2	Aav46370 Nucleic acid
3	575.4	83.4	1050	9	Adb58421 Toxicity
4	575.4	83.4	1050	9	Adb53005 Primary r
5	575.2	75.0	1059	3	Aaz38553 Human CXCR4
6	517.2	75.0	1059	6	Abq74938 Human CXCR4
7	517.2	75.0	1059	6	Aah99952 CXCR4 enc
8	517.2	75.0	1059	7	Acc2707 Human can
9	517.2	75.0	1225	6	Abz35630 Human sig
10	517.2	75.0	1225	7	ACA56637 Human sig
11	517.2	75.0	1317	2	Aav18357 Human RM3
12	517.2	75.0	1317	3	AAa91726 Human 7TM
13	517.2	75.0	1317	6	Abk4255 Human 7 t
14	517.2	75.0	1582	6	Abk4255 CXCR4 rec
15	517.2	75.0	1588	2	Aaz27610 Human CXCR4
16	517.2	75.0	1664	3	AAa61656 Human CXCR4
17	517.2	75.0	1664	3	Aaz40014 CXCR4 cod
18	517.2	75.0	1670	6	Abk3803 Human cDN
19	517.2	75.0	1670	6	Abn95645 Gene #214
20	517.2	75.0	1670	7	Acc45765 Human COP
21	517.2	75.0	1679	4	Aad12801 Human neu
22	517.2	75.0	1679	4	AAi65467 Nucleotid
23	517.2	75.0	1679	6	Abss3992 DNA encod

24	517.2	75.0	1679	6	Abss3750 DNA encod
25	517.2	75.0	1679	6	Abv78039 Hypoxia-r
26	517.2	75.0	1679	7	Abx08779 Angiogene
27	517.2	75.0	1679	7	Abx74454 Human cDN
28	517.2	75.0	1679	7	Abz68886 Nucleotid
29	517.2	75.0	1679	7	Abz42642 Human CXCR4
30	517.2	75.0	1679	9	Adc98645 Human CXCR4
31	517.2	75.0	1711	6	Abz35348 Human gen
32	517.2	75.0	1711	9	Adb47320 Human cDN
33	517.2	75.0	1737	2	AaQ80521 Human mon
34	517.2	75.0	1737	2	AAQ99007 Chemokine
35	517.2	75.0	1944	2	AAx15882 cDNA enco
36	517.2	75.0	1944	3	AAa34774 Human ade
37	517.2	75.0	5161	3	AAf20896 Human COR
38	517.2	75.0	5161	6	Aah99951 CXCR4 gen
39	517.2	75.0	5161	7	Abz26590 Human CCR
40	517.2	75.0	5161	7	ACA64750 Chemokine
41	515.6	74.7	1102	7	AAQ66179 Seven tra
42	515.6	74.7	1317	2	ADD67548 Human Lyl
43	515.6	74.7	8747	9	AAQ29506 New plate
44	514	74.5	1737	2	AAQ29506
45	512.4	74.3	1059	5	Abi97982 Non-endog

ALIGNMENTS

RESULT 1

AAZ27611

ID AAZ27611 standard; DNA; 1758 BP.

XX AC AAZ27611;

XX DT 16-DEC-1999 (first entry)

XX DE Mouse CXCR4 coding sequence.

XX CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;

KW tissue repairing agent; vascularisation; ss.

XX Mus sp.

XX WO9948528-A1.

XX PD 30-SEP-1999.

XX PF 23-MAR-1999; 99WO-JP001448.

XX PR 24-MAR-1998; 98JP-00095448.

XX (CHUS) CHUGAI SEIYAKU KK.

XX (KISH/) KISHIMOTO T.

XX Kishimoto T, Nagasawa T, Tachibana K;

XX WPI: 1999-591042/50.

XX P-PSDB; AAY39994.

XX CXCR4-potentiating agents and methods useful for inhibiting

XX neovascularization, and treating solid cancers.

XX Disclosure; page 50-51; 63pp; Japanese.

This sequence encodes the mouse CXCR4 protein. The invention relates to remedies inhibiting neovascularisation, remedies for solid cancer, remedies for diseases pathologically caused by neovascularisation and tissue repairing agents containing as the active ingredient a substance capable of potentiating CXCR4. Based on a finding that vascularisation is inhibited in a CXCR4 knockout mouse, it becomes possible to prepare remedies inhibiting vascularisation which contain as the active ingredient a substance capable of potentiating CXCR4, remedies for solid cancer, remedies for diseases pathologically caused by neovascularisation and tissue repairing agents containing as the active ingredient a

CC substance capable of potentiating CXCR4. It is also possible to establish
 CC methods for treatment with the use of these remedies
 XX
 SQ Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;
 Query Match 100.0%; Score 690; DB 2; Length 1758;
 Best Local Similarity 100.0%; Pred. No. 6.8e-186;
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCACACTCCCTCTTGGGCGAGTTGAT 60
 DB 238 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCACACTCCCTCTTGGGCGAGTTGAT 297
 QY 61 GCATGGGTGACTGTGACTTTGGGAAATTTTGTGAAGGTGTCATATCATCTACACT 120
 DB 298 GCATGGGTGACTGTGACTTTGGGAAATTTTGTGAAGGTGTCATATCATCTACACT 357
 QY 121 GTCACCTCTACAGCAGGTTCTCTCTGCTGCTTTCATCAGCTGACCGGTACCTGCC 180
 DB 358 GTCACCTCTACAGCAGGTTCTCTCTGCTGCTTTCATCAGCTGACCGGTACCTGCC 417
 QY 181 ATTGTCCAGCCCAACAGTCAAGGCAAGAACTGCTGGTGTGAAGGCGAGTCTAT 240
 DB 418 ATTGTCCAGCCCAACAGTCAAGGCAAGAACTGCTGGTGTGAAGGCGAGTCTAT 477
 QY 241 GTGGCGGTGATCCAGCCCTCTCTGACTATACCTGACTTCACTTTTGGCGAGTC 300
 DB 478 GTGGCGGTGATCCAGCCCTCTCTGACTATACCTGACTTCACTTTTGGCGAGTC 537
 QY 301 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTACCGCTTTACCCGAT 360
 DB 538 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTACCGCTTTACCCGAT 597
 QY 361 AGCTGTGGATGGTGGTGTTCATTCAGCATATAATGTGGGTCTCATCTCCCGCGC 420
 DB 598 AGCTGTGGATGGTGGTGTTCATTCAGCATATAATGTGGGTCTCATCTCCCGCGC 657
 QY 421 ATCGTCATCTCTCTGTTACTGATCATCATCTCTAAGCTGTACACATCCAGGGCAC 480
 DB 658 ATCGTCATCTCTCTGTTACTGATCATCATCTCTAAGCTGTACACATCCAGGGCAC 717
 QY 481 CAGAGCGCAAGCCCTCAAGACGACGTATCTCTCATCTCTAGCTTTCTTGGCTGTGG 540
 DB 718 CAGAGCGCAAGCCCTCAAGACGACGTATCTCTCATCTCTAGCTTTCTTGGCTGTGG 777
 QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCATCTCTTTTGGGAGTATCAAGCAA 600
 DB 778 CTGCCATATTATGTGGGATCAGCATCGACTCTCTCATCTCTTTTGGGAGTATCAAGCAA 837
 QY 601 GGATGTGACTTCGAGAGCATTTGTGACAAAGTGATCTCCATCAGAGGCGCTCGCCTTC 660
 DB 838 GGATGTGACTTCGAGAGCATTTGTGACAAAGTGATCTCCATCAGAGGCGCTCGCCTTC 690
 QY 661 TTCACCTGTGCTGAACCCCATCTCTAT 690
 DB 898 TTCACCTGTGCTGAACCCCATCTCTAT 927

RESULT 2
 AAV46370
 ID AAV46370 standard; cDNA to mRNA; 1877 BP.
 XX
 AC AAV46370;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Nucleic acid encoding a murine CXC chemokine receptor.
 XX
 KW Mouse; CXC chemokine receptor; pre-B cell line DM34;
 KW CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;
 KW screening; inhibitor; AIDS; ds.
 XX
 OS Mus sp.

XX Key Location/Qualifiers
 PH CDS 120..1199
 FT /*tag= a
 XX
 XX WO9835035-A1.
 PN 13-AUG-1998.
 PD 07-FEB-1997; 97WO-JP000299.
 PF 07-FEB-1997; 97WO-JP000299.
 PR (SHIO) SHIONOGI & CO LTD.
 XX
 XX Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
 PI Nakajima T, Yoshie O;
 XX
 DR WP1; 1998-447232/38.
 DR P-PSDS; AAW64778.
 XX
 PT Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
 PT factor is useful for screening of potential HIV infection and AIDS
 PT inhibitors.
 XX
 PS Claim 3; Page 39-42; 76pp; Japanese.
 XX
 CC The present sequence encodes a murine CXC chemokine receptor which binds
 CC to the mouse CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
 CC nucleic acid is isolated from mouse pre-B cell line DM34. The receptor
 CC and cells expressing it can be used in the study and mapping of the
 CC mechanism of HIV infection and in screening of potential inhibitors of
 CC HIV infection and the development of AIDS
 XX
 SQ Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;
 Query Match 100.0%; Score 690; DB 2; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 7e-186;
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTCTTGGCGAGTTGAT 60
 DB 357 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTCTTGGCGAGTTGAT 416
 QY 61 GCCATGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGGTGTCATATCATCTACACT 120
 DB 417 GCCATGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGGTGTCATATCATCTACACT 476
 QY 121 GTCACCTCTACAGCAGCGTTCTCTATCTCTGGCTTTCATCAGCTGACCGGTACCTCGCC 180
 DB 477 GTCACCTCTACAGCAGCGTTCTCTATCTCTGGCTTTCATCAGCTGACCGGTACCTCGCC 536
 QY 181 ATTGTCCAGCCCAACAGTCAAGGCAAGAACTGCTGGCTGAAAGGCGAGTCTAT 240
 DB 537 ATTGTCCAGCCCAACAGTCAAGGCAAGAACTGCTGGCTGAAAGGCGAGTCTAT 596
 QY 241 GTGGGGGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTCTATCTTTGCCGACGTC 300
 DB 597 GTGGGGGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTCTATCTTTGCCGACGTC 656
 QY 301 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
 DB 657 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 716
 QY 361 AGCTGTGGATGGTGGTGTTCATTTCCAGCATATAATGTGGGTCTCATCTCTGCCCGCC 420
 DB 717 AGCTGTGGATGGTGGTGTTCATTTCCAGCATATAATGTGGGTCTCATCTCTGCCCGCC 776
 QY 421 ATCGTCATCTCTCTGTTACTGTCATCATCTCTTAAGCTGTACACTTCCAGGGCGCAC 480
 DB 777 ATCGTCATCTCTCTGTTACTGTCATCATCTCTTAAGCTGTACACTTCCAGGGCGCAC 836
 QY 481 CAGAGCGCAAGCCCTCAAGACGACGTATCTCTCATCTCTAGCTTTCTTTGCCCTGCTGG 540

Db 837 CAGAAAGGCGAGGCGCTCAAGACGACGATCATCTCATCTAGCTTTCTTTGGCTGCTGG 896
Qy 541 CTGCCATATTATGTGGGATCAGATCGACTCCTTCATCCTTTTGGGAGTCAACGAA 600
Db 897 CTGCCATATTATGTGGGATCAGATCGACTCCTTCATCCTTTTGGGAGTCAACGAA 956
Qy 601 GGATGTGACTTTCGAGAGATGTGCACAGTGGATCTCCATCAGAGGCGCTCGCCTTC 660
Db 957 GGATGTGACTTTCGAGAGATGTGCACAGTGGATCTCCATCAGAGGCGCTCGCCTTC 1016
Qy 661 TTCCACTGTTGCTGAACCCCATCTCTAT 690
Db 1017 TTCCACTGTTGCTGAACCCCATCTCTAT 1046

RESULT 3

ADBS58421
ID ADB58421 standard; DNA; 1050 BP.

AC ADB58421;

DT 04-DEC-2003 (first entry)

XX Toxicity-related gene, SEQ ID 3447.

DE Toxic; toxin; gene expression profile; hepatotoxicity; liver;

KW drug screening; toxicity assay; ds.

XX Unidentified.

OS Unidentified.

XX WO2003064624-A2.

PN 07-AUG-2003.

XX 31-JAN-2003; 2003WO-US003194.

XX 31-JAN-2002; 2002US-00060087.

PR 15-MAR-2002; 2002US-0364045P.

PR 15-MAR-2002; 2002US-0364055P.

PR 30-DEC-2002; 2002US-0436643P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Potter M, Johnson K, Higgs B, Castle A, Elashoff M;

PI WPI; 2003-689530/65.

XX Predicting a toxic effect of a compound, useful in identifying toxicity

XX markers in liver tissues or cells for drug screening and toxicity assays,

XX comprises preparing gene expression profile of tissue or cells exposed to

XX the compound.

XX Claim 1; SEQ ID NO 3447; 1156pp; English.

XX The present invention relates to a method for predicting a toxic effect

XX of a compound. The method comprises preparing a gene expression profile

XX of a tissue or cell sample exposed to the compound, and comparing the

XX gene expression profile to a database comprising SEQ ID 1-4925, where

XX differential expression of the gene indicates at least one toxic effect.

XX The method is useful for predicting at least one toxic effect of a

XX compound, predicting hepatotoxicity or the progression of a toxic effect

XX of a compound, identifying an agent that modulates the onset or

XX progression of a toxic response, predicting the cellular pathways that a

XX compound modulates in a cell, and identifying an agent that modulates at

XX least one activity of a protein. The method and compositions of the

XX present invention using a database of genes having liver toxin-induced

XX differential expression, are useful in identifying toxicity markers in

XX liver tissues or cells for drug screening and toxicity assays. Note: The

XX sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1050 BP; 231 A; 313 C; 247 G; 259 T; 0 U; 0 Other;

Query Match 83.4%; Score 575.4; DB 9; Length 1050;

Best Local Similarity 91.1%; Pred. No. 2.3e-153;

Matches 628; Conservative 0; Mismatches 46; Indels 15; Gaps 1;

Qy 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTGCATCAGCTCCCTCTCTGGGAGTTGAT 60

Db 223 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTGCATCAGCTCCCTCTCTGGGAGTTGAC 282

Qy 61 GCATGGCTGACTGTGTACTTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACAT 120

Db 283 GCATGGCTGACTGTGTACTTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACAC 342

Qy 121 GTCAACCTCTACAGCAGGCTTCTCATCTCTGGCTTCTCATCAGCTCGACGGTACCTCGCC 180

Db 343 GTCAACCTTTACAGCAGGCTTCTCATCTCTGGCTTCTCATCAGCTCGACGGTACCTCGCC 402

Qy 181 ATTGTCCACGCCCAACCAAGTCAAGGCCAAGGAACTGTGGCTGAAAGGAGCTAT 240

Db 403 ATTGTCCACGCCCAACCAAGTCAAGGCCAAGGAACTGTGGCTGAAAGGAGCTAT 462

Qy 241 GTGGGCTGTGATCCAGCCCTCTCTCTGACTATACCTGACTTCTATCTTTCGCGAGCTC 300

Db 463 GTGGGCTGTGATCCAGCCCTCTCTCTGACTATACCTGACTTCTATCTTTCGCGAGCTC 522

Qy 301 AGCCAGGGGACATCAGTCCAGGGGATGACAGGTATCATCTGTGACCGCTTTACCCCGAT 360

Db 523 AGCCAGGGGAC-----GGCAGGTATCATCTGTGACCGCTTTACCCCGAC 567

Qy 361 AGCTGTGATGGTGGTGTTCATTTCCAGATATAAATGGTGGGTCTCATCTTTCGCGAGCTC 420

Db 568 AGCTGTGATGGTGGTGTTCATTTCCAGATATAAATGGTGGGTCTCATCTTTCGCGAGCTC 627

Qy 421 ATCGTCATCTCTCTCTGTGTACTGTCATCATCTCTTAAGCTGTACATCTCCAGAGGCTCAC 480

Db 628 ATCGTCATCTCTCTGTGTACTGTCATCATCTCTTAAGCTGTACATCTCCAGAGGCTCAC 687

Qy 481 CAGAGCGCAGCGCCCTCAAGACGACGATCATCTCTCAAGCTGTACATCTCCAGAGGCTCAC 540

Db 688 CAGAGCGCAGCGCCCTCAAGACGACGATCATCTCTCAAGCTGTACATCTCCAGAGGCTCAC 747

Qy 541 CTGCCATATTATGTGGGATCAGATCGACTCCTCTTCATCTTTTGGGAGTCAACGAA 600

Db 748 CTGCCATATTATGTGGGATCAGATCGACTCCTCTTCATCTTTTGGGAGTCAACGAA 807

Qy 601 GGATGTGACTTTCGAGAGCATTTGTGCACAGTGGATCTCCATCAGAGAGCCCTCGCCTTC 660

Db 808 GGATGTGACTTTCGAGAGCATTTGTGCACAGTGGATCTCCATCAGAGAGCCCTCGCCTTC 867

Qy 661 TTCCACTGTTGCTGAACCCCATCTCTA 689

Db 868 TTCCACTGTTGCTGAACCCCATCTCTA 896

RESULT 4

ADBS53005
ID ADB53005 standard; DNA; 1050 BP.

XX ADB53005;

XX 04-DEC-2003 (first entry)

XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3547.

XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;

XX toxicity marker; toxicity progression; drug screening;

XX primary rat hepatocyte toxicity modelling; gene; ds.

XX Rattus norvegicus.

XX WO2003065993-A2.

XX

XX

PI Murphy GP, Boynton AL, Sehgal A;
 XX WPI; 2000-052634/04.
 DR P-PSDB; AAY52507.
 XX
 PT Use of CXCR-4 and SDF-1 as markers for diagnosis and treatment of e.g.
 PT tumors, degenerative disorders, growth deficiencies, hyper- and
 PT hypoproliferative disorders, physical trauma, lesions and wounds.
 XX
 PS Disclosure; Fig 14; 138pp; English.
 XX
 CC This sequence represents cDNA encoding CXCR-4 (CXCR-4).
 CC Chemokine receptors play an important role in the chemotaxis of T-cells
 CC and phagocytic cells to areas of inflammation. CXCR-4 is a member of the
 CC G-protein-coupled receptor family, which is involved in signal
 CC transduction, and its ligand is stromal cell derived factor-1 (SDF-1,
 CC AAY52508). CXCR-4 also mediates CD4-independent infection by HIV-1. CXCR-
 CC 4 has now been found to have a role in the aberrant proliferative
 CC behaviour of a number of cell types, including numerous primary tumours
 CC and derived cell lines. CXCR-4 is involved in cell transformation and
 CC tumorigenesis, particularly in brain, breast and colon tumours. It was
 CC found to be overexpressed in several brain tumour derived cell lines and
 CC primary brain tumour tissues, breast tumour tissues, colorectal
 CC adenocarcinoma, lung carcinoma and melanoma cell lines. CXCR-4 expression
 CC was required for the continuous proliferation of glioblastoma cancer
 CC cells, and inhibition of its gene function resulted in growth arrest.
 CC Conversely, overexpression resulted in enhanced and rapid cellular
 CC proliferation. CXCR-4 and SDF-1 can be used as markers for the diagnosis
 CC and prognosis of cell proliferative disorders, and antisense
 CC oligonucleotides complementary to at least a portion of an RNA transcript
 CC of a CXCR-4 gene can be used to inhibit hyperproliferation of a tumour
 CC cell. Agents that inhibit CXCR-4 function can be used for treating or
 CC preventing a disease or disorder involving cell overproliferation, e.g.,
 CC brain cancer, breast cancer, colon cancer, prostate cancer and B cell
 CC lymphoma, and also preneoplastic conditions, benign tumours
 CC hyperproliferative disorders, and benign dysproliferative disorders. They
 CC can also be used for treating e.g., cirrhosis of the liver, keloid
 CC formation, psoriasis, benign tumors, fibrocystic conditions and tissue
 CC hypertrophy. Compounds that promote CXCR-4 function can also be used for
 CC preventing or treating a disease or disorder involving a deficiency in
 CC cell proliferation, or treating a condition where cell proliferation
 CC would be desirable. Such diseases include degenerative disorders (e.g.,
 CC Parkinson's disease, Alzheimer's disease), growth deficiencies,
 CC hypoproliferative disorders, physical trauma, lesions (e.g., those caused
 CC by ischaemia), and wounds
 XX
 SQ Sequence 1059 BP; 244 A; 301 C; 232 G; 282 T; 0 U; 0 Other;
 Query Match 75.0%; Score 517.2; DB 3; Length 1059;
 Best Local Similarity 85.8%; Pred. No. 8.8e-137;
 Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
 QY 1 CTGCACCTGTGACGCTGACCTCTCTTTGTCATCACACTCCCTCTCTGGCAGTTGAT 60
 DB 232 CTGCACCTGTGACGCTGACCTCTCTTTGTCATCACACTCCCTCTCTGGCAGTTGAT 291
 QY 61 GCCATGGCTGACTGTTCTGGGAAATTTTGTGTAAGGTGTCATATCACTACT 120
 DB 292 GCCGTGGCAAACTGGTACTTTGGGAACCTTCATGCAAGGCAGTCCATGTCATCACACA 351
 QY 121 GTCACCTCTACAGCAGGTTCTCATCTGGCCCTTCATCAGCTGGACCGTACTCGCC 180
 DB 352 GTCACCTCTACAGCAGTGTCTCTATCTGGCCCTTCATGTCAGTGGACCGTACTCGCC 411
 QY 181 ATTGTCCACGCCAACACAGTCAAGGCCAAGGAACATGCTGGCTGAAAGGCGAGTCTAT 240
 DB 412 ATCGTCCACGCCAACACAGTCAAGGCCAAGGAAGCTGTTGGCTGAAAGGCTGCTAT 471
 QY 241 GTGGGCGTCTGGATCCCAAGCCCTCTCTCTGACTATATCTGACTTCATCTTTGGCAGTGC 300
 DB 472 GTTGGCGTCTGGATCCCTGCGCCCTCTCTGCTGACTATATCCCGACTTCATCTTTGCCAACG-- 529
 QY 301 AGCCAGGGGGACATCAGTCAGGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360

DB 530 -----TCAGTGAGGCGAGATGACAGATATATCTGTGACCGCTTTTACCCCAAT 576
 QY 361 AGCCTGTGGATGGTGGTGTCTTCAATTCAGCATATAATAGTGGGTCTCATCTGCGCGGC 420
 DB 577 GACTTGTGGTGGTGTGTGTTCCAGTTTCAGCACATCATGGTTGGCTTATCTGCTGCTGT 636
 QY 421 ATCGTCACTCTCTCTGTTACTGCAATCATCTCTTAAGCTGTGACATCCAGGCGCAC 480
 DB 637 ATTGTCACTCTCTCTGTTACTGCAATCATCTCTTAAGCTGTGACATCCAGGCGCAC 696
 QY 481 CAGAGCGCAGAGCCCTCAAGACGACAGTCACTCTCTATCTCTTCTTGGCTGTGCG 540
 DB 697 CAGAGCGCAGAGCCCTCAAGACGACAGTCACTCTCTATCTCTGCTGTGCG 756
 QY 541 CTGCCATATATATGGGATGAGCATGCACTCTCTTCACTCTTTTGGAGTATCAAGCAA 600
 DB 757 CTGCTTACTTACTATGGGATGAGCATGCACTCTCTTCACTCTTCTCTGAAATCATCAAGCAA 816
 QY 601 GGATGTGACTTCCAGAGCATTTGTGCACAAGTGGATCTCCATCACAAGGCGCTTCCCTTTC 660
 DB 817 GGGTGTGAGTTTGAGAACACTGTGCAAGTGGATTTCATCACCAGGCGCTTCTTTC 876
 QY 661 TTCCACTGTGTGCTGAACCCCATCTCTAT 690
 DB 877 TTCCACTGTGTGCTGAACCCCATCTCTAT 906
 RESULT 6
 ABQ74938
 ID ABQ74938 standard; cDNA; 1059 BP.
 XX
 AC ABQ74938;
 XX
 DT 30-OCT-2002 (first entry)
 XX
 DE Human CXCR chemokine receptor 4 encoding cDNA SEQ ID NO:125.
 XX
 KW Cysteine-X-cysteine chemokine receptor 4; CXCR chemokine receptor 4;
 KM binding; identification; AIDS; HIV; CXCR4; virucide; anti-HIV; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1059
 FT CDS /tag= a
 FT /product= "Cysteine-X-cysteine chemokine receptor 4"
 XX
 PN WO200257313-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 26-OCT-2001; 2001WO-US051165.
 XX
 PR 27-OCT-2000; 2000US-0243587P.
 PR 20-MAR-2001; 2001US-00813448.
 PR 20-MAR-2001; 2001US-00813651.
 PR 20-MAR-2001; 2001US-00813653.
 XX
 PA (CONS-) CONSENSUS PHARM INC.
 XX
 PI Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA, Krstenansky J;
 XX
 DR WPI; 2002-608378/65.
 DR P-PSDB; ABP52851.
 XX
 PT Identifying a binding compound for cysteine-X-cysteine chemokine receptor
 PT 4 (CXCR4) for preventing or treating AIDS and HIV infection, comprises
 PT binding a molecule to a molecule with a binding property corresponding to
 PT CXCR4.
 XX
 PS Example 3; Fig 3A-B; 92pp; English.
 XX

QY 121 GTCAACCTCTACAGCAGGTTCTCATCTCTGGCCTTCATCAGCTGACCGGTACCTCGCC 180
Db 352 GTCAACCTCTACAGCAGGTTCTCATCTCTGGCCTTCATCAGCTGACCGGTACCTCGCC 411
QY 181 ATTGTCCACGCCCAACCAAGCTCAAGGCAAGGAACTGCTGGCTGAAAGGAGGCTCTAT 240
Db 412 ATCGTCCACGCCCAACCAAGCTCAAGGCAAGGAACTGCTGGCTGAAAGGAGGCTCTAT 471
QY 241 GTGGGCTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCTCTTTGGCCGAGCTC 300
Db 472 GTTGGGCTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCTCTTTGGCCGAGCTC 529
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTACCGCCCTTTACCCCGAT 360
Db 530 -----TCAGTGAGGACATGACAGATATATCTGTGACCGCTTTACCCCAAT 576
QY 361 AGCCTGTGGATGGTGTTCATTCAGCATATAATGGTGGGTCTCATCTCTGCGGCG 420
Db 577 GACTTGTGGGTTGTGTTCAGTTTCAGCATATCATGTGTGGCTTATCTGCTGCT 636
QY 421 ATCGTCACTCTCTCTGTACTGATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 637 ATTGTCACTCTCTCTGTCT 696
QY 481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 697 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 756
QY 541 CTGGCATATATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 757 CTGGCATATATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 816
QY 601 GGATGTGACTTCAGAGCATTGTGCAAGTGTGATCTCCATCAGAGGCGCTCGGCTTC 660
Db 817 GGGTGTGAGTTTGAACACATGTGCAAGTGTGATTTCCATCAGGAGCGCTAGCTTTC 876
QY 661 TTCCACTGTGTGCTGAACCCCATCTCTAT 690
Db 877 TTCCACTGTGTGCTGAACCCCATCTCTAT 906
RESULT 8
ACC72707
ID ACC72707 standard; cDNA; 1059 BP.
XX
AC ACC72707;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein encoding cDNA SEQ ID NO:46.
XX
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350666P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI

PI Zlotnik A;
XX
DR WPI; 2003-354600/33.
DR P-PSDB; ABR58580.
XX
PT New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
XX Claim 8; Page 650; 767pp; English.
XX
XX The present invention describes an isolated nucleic acid molecule, which
XX comprises the sequence of any of the genes that are up-regulated or down-
XX regulated in specific cancers (e.g. about 1031 genes up-regulated in
XX acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
XX related gene nucleotide sequences which encode the proteins given in
XX ABR5851 to ABR5870. Also described: (1) determining the presence or
XX absence of a pathological cell in a patient; (2) an expression vector
XX comprising a nucleic acid molecule described above; (3) a host cell
XX comprising the vector; (4) an isolated polypeptide, which is encoded by
XX the nucleic acid; (5) an antibody that specifically binds the polypeptide
XX of (4); (6) specifically targeting a compound to a pathological cell in a
XX patient by administering to the patient the antibody above; and (7) a
XX drug screening assay. The nucleic acid is useful as diagnostic markers or
XX therapeutic targets. In particular, the nucleic acid is useful for
XX diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
XX bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
XX pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
XX atherosclerosis and endometriosis. The nucleic acid is also useful in
XX drug screening, particularly for identifying agents for treating these
XX pathologies
SQ Sequence 1059 BP; 244 A; 300 C; 232 G; 283 T; 0 U; 0 Other;
Query Match 75.0%; Score 517.2; DB 7; Length 1059;
Best Local Similarity 85.8%; Pred. No. 8.8e-137;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGCACCTGTGAGTGGGTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 232 CTGCACCTGTGAGTGGGTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291
QY 61 GCCATGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACT 120
Db 292 GCCGTGGCAAACTGTGACTTTGGGAACTTCTATGCAAGGCGAGTCCATGTATCATACACA 351
QY 121 GTCAACCTCTACAGCAGGTTCTCATCTCTGGCCTTCATCAGCTTGGACCGGTACCTCGCC 180
Db 352 GTCAACCTCTACAGCAGGTTCTCATCTCTGGCCTTCATCAGCTTGGACCGGTACCTCGCC 411
QY 181 ATTGTCCAGGCCCAACCAAGCTCAAGGCAAGGAACTGCTGGCTGAAAGGAGGCTCTAT 240
Db 412 ATCGTCCACGCCCAACCAAGCTCAAGGCAAGGAACTGCTGGCTGAAAGGAGGCTCTAT 471
QY 241 GTGGGCTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCTCTTTGGCCGAGCTC 300
Db 472 GTTGGGCTCTGGATCCAGCGCTCTCTCTGACTATTCGAGCTTCTCTTTGGCCGAGCTC 529
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
Db 530 -----TCAGTGAGGACATGACAGATATATCTGTGACCGCTTTACCCCAAT 576
QY 361 AGCCTGTGGATGGTGTTCATTCAGCATATAATGGTGGGTCTCATCTCTGCGGCG 420
Db 577 GACTTGTGGGTTGTGTTCAGTTTCAGCATATCATGTGTGGCTTATCTGCTGCT 636
QY 421 ATCGTCACTCTCTCTGTACTGATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 637 ATTGTCACTCTCTGTCT 696
QY 481 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 697 CAGAAGCGCAAGGCCCTCAAGACAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 756

QY	541	CTGCCATATTATGTGGGATCAGCATGCATCTCTTCTCATCTCTTTTGGAGTTCATCAAGCA	500
Db	757	CTGCGCTTACTACATTGGGATCAGCATGCATCTCTTCTCATCTCTTCTGAAATCATCAAGCA	816
QY	601	GGATGTGACTTTCGAGAGCAATTGTGCACAAGTGGATCTCCATCACAGAGCCCTTCGCTTC	660
Db	817	GGGTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCATCACAGGAGCCCTAGCTTTC	876
QY	661	TTCCACTGTTTGCCTGAACCCCATCTCTAT	690
Db	877	TTCCACTGTTTGTCTGAACCCCATCTCTAT	906

RESULT 9
 AB235630
 IID AB235630 standard; cDNA; 1225 BP.
 XX
 AC
 XX
 AC
 XX
 XX
 05-FEB-2003 (first entry)
 XX
 Human gene expression profile polynucleotide SEQ ID NO 741.
 XX
 Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 KW
 KW
 OS
 Homo sapiens.
 PN
 WO200274979-A2.
 PD
 26-SEP-2002.
 PF
 20-MAR-2002; 2002WO-008456.
 PP
 20-MAR-2001; 2001US-0276947P.
 PA
 (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX
 Wan J, Wang Y;
 FI
 WPI; 2002-740862/80.
 XX
 New gene expression profile generated from primary, endothelial,
 PT epithelial, and muscle cell types, useful for identifying disease
 PT pathologies involving alterations of gene expression, e.g. cancer.
 XX
 Example 3; Page 827-828; 850pp; English.
 XX
 The invention relates to a gene expression profile comprising one or more
 CC genes (AB234889-AB235692) and generated from a cell type. The cell type
 CC is a coronary artery, endothelium, umbilical artery or vein endothelium,
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery,
 CC endothelium, myometrium microvascular endothelium, keratinocyte
 CC epithelium, bronchial epithelium, mammary epithelium, prostate
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
 CC small airway epithelium, renal epithelium, umbilical artery smooth
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast.
 CC osteoblasts or prostate stromal cell. The gene expression profile is used
 CC for determining the level of RNA expression for a sample, determining the
 CC phenotype of a cell and distinguishing cell types. The gene or a protein
 CC expression profile is useful in identifying disease pathologies involving
 CC alterations of gene expression. The assessment of expression profiles may
 CC provide meaningful information with respect to tumour type and stage, the
 CC treatment methods, and prognosis. The gene or protein expression profile
 CC may also be used for creating microarrays. The microarray is useful for
 CC genetic and physical mapping of genomes. DNA sequencing, genetic or
 CC medical diagnosis, genotyping of organisms, confirming cellor tissue

XX US6500938-B1.
XX 31-DEC-2002.
XX 30-JAN-1998; 98US-00016434.
XX 30-JAN-1998; 98US-00016434.
XX (INCY-) INCYTE GENOMICS INC.
XX Au-Young J, Seilhamer JU;
XX WPI; 2003-352189/33.
XX
XX Combination of polynucleotide probes, useful as array elements in a
XX microarray for monitoring the expression of a number of target
XX polynucleotides.
XX Claim 1; SEQ ID NO 1235; 65pp; English.
XX
XX The invention relates to a combination which, comprises a number of
XX polynucleotide probes comprising a sequence selected from one of the 1490
XX sequences mentioned in the specification. The combination is useful as an
XX array element in a microarray for monitoring the expression of a number
XX of target polynucleotides. The microarray is particularly useful in the
XX diagnosis and treatment of cancer and immunopathology and neuropathology.
XX The microarray is useful in diagnostics and treatment regimens, drug
XX discovery and development, toxicological and carcinogenicity studies,
XX forensics and pharmacogenomics. The microarray is also useful for
XX monitoring progression of diseases and for developing sophisticated
XX profiles for the effects of currently available therapeutic drugs. The
XX combination is also useful for purifying a subpopulation of mRNAs, cDNAs
XX and genomic fragments and in research and diagnostic applications. The
XX array can detect changes in expression in a large number of genes coding
XX for different signaling pathway populations which can be used to diagnose
XX various diseases including cancer e.g. adenocarcinoma and leukaemia,
XX immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
XX and Parkinson's disease. The present sequence represents a polynucleotide
XX probe of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docID=06500938B1
XX
XX Sequence 1225 BP; 288 A; 324 C; 259 G; 354 T; 0 U; 0 Other;
XX
XX Query Match 75.0%; Score 517.2; DB 7; Length 1225;
XX Best Local Similarity 85.8%; Pred. No. 9.3e-137;
XX Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
XX
Qy 1 CTGACCTGTGACGTGACCTCTCTTTGTGTCATCAGCTCCCTCTCTGGGAGTTGAT 60
Db |||||
257 CTGACCTGTGACGTGACCTCTCTCTTTGTGTCATCAGCTCCCTCTCTGGGAGTTGAT 316
Qy |||||
61 GCCATGGCTGACGTGACCTCTCTTTGGAAATTTTCTGTAAGGCTGTCCATATCATCTACCT 120
Db |||||
317 GCGTGCAAACTGGTACTTTGGAACTTCTTAAGGAGCTGCTATCATCTACACA 376
Qy |||||
121 GTCAACCTCTACAGCAGCTTCTCATCTGGCTTTCATCAGCTGGACCGGTACTCGCC 180
Db |||||
377 GTCAACCTCTACAGCAGCTGCTCTCATCTGGCTTTCATCAGCTGGACCGGTACTCGCC 436
Qy |||||
181 ATTGTCCACGCCACCACTGTAAGGCCAAGAACTGCTGCTGAAAGGAGCTCTAT 240
Db |||||
437 ATGCTCCACGCCACCACTGTAAGGCCAAGAACTGCTGCTGAAAGGAGCTCTAT 496
Qy |||||
241 GTGGGCTCTGGATCCAGCCCTCTCTGACTATATACCTGACTTTCATCTTTGGCCAGCTC 300
Db |||||
497 GTTGGCTCTGGATCCCTGCTCTCTGCTGCTATTTCCGACTTTCATCTTTGCCAAG-- 554
Qy |||||
301 AGCCAGGGGAGCATCAGTCAAGGGGAGTACAGGTACATCTGTGACCGCTTTACCCCGAT 360
Db |||||
555 -----TCAGTGAGGAGAGTACAGATATATCTGTGACCGCTTCTACCCCAAT 601

Qy 361 AGCCTGTGGATGGTGGTGTCTTCAATTCAGCATATATATGGTGGTCTCATCTCTCCGGC 420
Db |||||
602 GACTTGTGGTGGTGGTGTCTTCCAGTTTCAGCATCATATGGTGGTGGTGTCTCTGCTGT 661
Qy 421 ATCGTCATCTCTCTCTGTACTGTGATCATCTCTAAGCTGTGACATCTCAAGGGCCAC 480
Db |||||
662 ATTGTCTATCTCTCTGTCTATTGCTATCATCTCCAGCTGTGACATCTCAAGGGCCAC 721
Qy 481 CAGAAGCGCAAGGGCCCTCAAGACGACAGTCTATCTCTATCTCTCTCTCTCTCTCTCT 540
Db |||||
722 CAGAAGCGCAAGGGCCCTCAAGACGACAGTCTATCTCTCTCTCTCTCTCTCTCTCTCT 781
Qy 541 CTGCCATATTATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db |||||
782 CTGCTTACTACATTTGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 841
Qy 601 GGATGTGACTTTCGAGAGCATTTGTGCAAGTGGATCTCCATCAGAGGCGCTCTGCTTC 660
Db |||||
842 GGGTGTGAGTTTGAGAACACATGTGCAAGTGGATTTCCATCAGCGAGGCGCTAGCTTTC 901
Qy 661 TTCCACTGTTGCTGCTGAAACCCCATCTCTCTAT 690
Db 902 TTCCACTGTTGCTGAAACCCCATCTCTCTAT 931
XX
XX AAV18357 standard; DNA; 1317 BP.
XX AAV18357
XX AC AAV18357;
XX DT 25-SEP-1998 (first entry)
XX DE Human RM3 seven transmembrane (7TM) receptor cDNA.
XX KW V28; placenta; seven transmembrane receptor; 7TM; signal transduction;
XX immunology; inflammation; RM3; 88.
XX OS Homo sapiens.
XX FH Key
XX CDS 201..1211
XX /*tag= a
XX /product= "Human RM3 seven transmembrane receptor"
XX
XX US5759804-A.
XX PN 02-JUN-1998.
XX PD 17-NOV-1993; 93US-00153848.
XX PF 17-NOV-1992; 92US-00977452.
XX PR (ICOS-) ICOS CORP.
XX PA
XX PI Schweickart VL, Godiska R, Gray PW;
XX WPI; 1998-332132/29.
XX DR P-PSDB; AAW48734.
XX DX
XX PT DNA encoding V28 seven transmembrane receptor polypeptide - useful for
XX producing recombinant polypeptide and anti-V28 antibodies, and in
XX screening assays for V28 agonists and antagonists.
XX PS Example 11; Col 89-94; 56pp; English.
XX
XX The present novel sequence represents the human RM3 cDNA isolated from a
XX human macrophage cDNA library. The invention claims for the full length
XX V28 genomic DNA sequence (AAV18343) isolated from a human placenta
XX genomic library. The V28 (AAW48722) and RM3 proteins are seven
XX transmembrane (7TM) receptors which are probably involved in signal
XX transduction. The invention also claims that cells transformed with V28

DT	15-NOV-2002	(first entry)
XX	CXCR4	receptor DNA sequence.
XX	Receptor;	virus infection; immune response; autoimmune disease;
KW	TNFR1	receptor; inflammation; autoimmune encephalomyelitis;
KW	rheumatoid	arthritis; myocarditis; CXCR4 receptor; ds.
XX	Unidentified.	
OS	WO200257308-A2.	
PN	25-JUL-2002.	
PD	22-JAN-2002;	2002WO-GB000246.
XX	22-JAN-2001;	2001GB-00001576.
XX	07-FEB-2001;	2001GB-00003032.
PR	(SANG-)	SANGAMO BIOSCIENCES INC.
PA	Moore M, Isalan M, Reynolds L, Ullman C, Girdlestone J;	
PI	Demaison C, Choo Y;	
XX	WPI;	2002-590720/63.
DR	New polypeptide	capable of binding nucleic acids for treating or
XX	preventing a disease	caused by a virus or of a disease associated with an
PT	immune response,	particularly autoimmune diseases e.g. inflammation or
PT	rheumatoid arthritis.	
XX	Disclosure;	Page 12-13; 105pp; English.
XX	The present invention	relates to a new polypeptide capable of binding to
CC	a nucleic acid	comprising a receptor nucleotide sequence which can
CC	function as a	receptor for virus infection or is involved in an immune
CC	response. The	molecules of the invention are useful in the treatment or
CC	prevention of a	disease caused by a virus or of a disease associated with
CC	an immune	response, particularly an autoimmune disease. The zinc finger
CC	polypeptide or a	nucleic acid encoding such a polypeptide can be used to
CC	modulate	transcription of a receptor nucleotide sequence. The nucleic
CC	acid polypeptide	capable of binding to a nucleic acid sequence comprising
CC	a receptor	nucleotide sequence, where the receptor is capable of
CC	functioning as a	receptor for infection by the virus or is involved in an
CC	immune	response, or a nucleic acid encoding such a polypeptide, can be
CC	used in the	preparation of a medicament for use in the treatment or
CC	prevention of a	disease caused by a virus or of a disease associated with
CC	an immune	response. The polypeptides may further be used to treat or
CC	prevent various	diseases or syndromes associated with or caused by
CC	malfunction of	the receptor as the TNFR1 receptor, such as autoimmune
CC	diseases	including inflammation, autoimmune encephalomyelitis, rheumatoid
CC	arthritis or	myocarditis. The present nucleic acid sequence represents a
CC	CXCR4	receptor sequence, as described in the invention
XX	Sequence	1582 BP; 386 A; 376 C; 333 G; 487 T; 0 U; 0 Other;
QQ	Query Match	75.0%; Score 517.2; DB 6; Length 1582;
XX	Best Local Similarity	85.8%; Pred. No. 1e-136;
XX	Matches 592;	Conservative 0; Mismatches 83; Indels 15; Gaps 1
QY	1	CTGCACCTGTCAGTGGCTGACCTCCTCTTTGTCATCACATCCCTCTCGGGCAGTTGAT 60
DB	223	CTGCACCTGTCAGTGGCGACCTCCTCTTTGTCATCACGTTTCCCTTCTGGGCGATTGAT 282
QY	61	GCCATGGCTGACTGGTACTTTTGGGAAATTTTGTGTAAAGCTGTCCATATCATCTACACT 120
DB	283	GCGGTGGCAAACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCTATCACACA 342
QY	121	GTCAACCTCTACAGCAGCGTTCTCATCTCGGCTTTCATCAGCTGGACCGGTACCTTCGCC 180
DB	343	GTCAACCTCTACAGCAGTGTCTCATCTCTGGCTTTCATCAGTCTGGACCGGTACCTTCGCC 402
QY	181	ATTGTCCAGCCCAACACAGTCAAAAGGCCAAGGAAACTGTGGCTCGAAAGGCAGTCTAT 240

CC remedies inhibiting neovascularisation, remedies for solid cancer,
CC remedies for diseases pathologically caused by neovascularisation and
CC tissue repairing agents containing as the active ingredient a substance
CC capable of potentiating CXCR4. Based on a finding that vascularisation is
CC inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC remedies inhibiting vascularisation which contain as the active
CC ingredient a substance capable of potentiating CXCR4, remedies for solid
CC cancer, remedies for diseases pathologically caused by neovascularisation
CC and tissue repairing agents containing as the active ingredient a
CC substance capable of potentiating CXCR4. It is also possible to establish
CC methods for treatment with the use of these remedies
XX
SQ Sequence 1588 BP; 396 A; 375 C; 334 G; 483 T; 0 U; 0 Other;

Query Match		75.0%; Score 517.2; DB 2; Length 1588;
Best Local Similarity		85.8%; Pred. No. 1e-136;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;		
Qy	1	CTGCACCTGTCAGTGGCTGACCTCTCTTTGTATACACTCCCTCTCGGGCAGTTGAT 60
Db	232	CTGCACCTGTCAGTGGCGGACCTCTCTTTGTATACAGCTTCCCTCTCGGGCAGTTGAT 291
Qy	61	GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAAGCTGTCCATATCATCTACACT 120
Db	292	GCCGTGGCAAACTGGTACTTTGGGAAGCTTCTATGCAAGGCAGTCCATGTCTATACACA 351
Qy	121	GTCAACCTCTACAGACAGGTTCTCATCTGGCCCTTCATCAGCCTGGACCGGTACCTCGCC 180
Db	352	GTCAACCTCTACAGACAGTGCCTCATCTGGCCCTTCATCAGTGTGACCGCTACCTGGCC 411
Qy	181	ATTGTCCACGCCCAACACAGTCAAAAGGCCAAGGAAAGCTGTGGCTGAAAGGCAGTCTAT 240
Db	412	ATCGTCCACGCCCAACACAGTCAAGAGGCCAAGGAGCTGTGGCTGAAAGGTGGTCTAT 471
Qy	241	GTGGGCGTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTGCCGACGTC 300
Db	472	GTGGGCGTCTGGATCCCTGCGCTCTCTCTGACTATTCCTCGACTTCATCTTTGCCAAGC-- 529
Qy	301	AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
Db	530	-----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 576
Qy	361	AGCCTGTGGATGGTGGTTTCAATTCACAGCATATATATGGTGGGTCTCATCTCTGCCCGC 420
Db	577	GACTTGTGGGTGGTGGTTCCAGTTTCAGCACATCATGGTTGGCCTTATCTCTGCCCTGGT 636
Qy	421	ATCGTCATCTCTCTCTTACTGCTCATCATCTCTTAAGCTGTACACTCCAGGGGCCAC 480
Db	637	ATTGTCACTCTCTCTCTTATGATATATCTTCAAGCTGTACACTCCAGGGGCCAC 696
Qy	481	CAGAAGCGCAAGGCCCTCAAGACAGCAGTCACTCTCATCTCTAGCTTTCTTTGGCTGCTGG 540
Db	697	CAGAAGCGCAAGGCCCTCAAGACACAGTCACTCTCATCTCTGGCTTTCTTGGCTGTTGG 756
Qy	541	CTGCCATATTATGTGGGGATCAGATCGACTCTTTCATCTCTTTTGGGAGTCAAGCAA 600
Db	757	CTGCCCTTACTACATTTGGGATCAGATCGACTCTTTCATCTCTCTGGAAATCATCAAGCAA 816
Qy	601	GGATGTCACTTCAGACATTTGTGCAAGTGGATCTCCATCAGAGGGCCCTCGCCTTC 660
Db	817	GGGTGTGAGTTTGAACACCTGTGCAAGTGGATTTCCATCAGAGGGCCCTAGCTTTC 876
Qy	661	TTCCACTGTTGGCTGAACCCCACTCTCTAT 690
Db	877	TTCCACTGTTGCTGAACCCCACTCTCTAT 906

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 55.2446 Seconds
 (without alignments)
 6931.281 Million cell updates/sec

Title: US-09-367-052-3
 Perfect score: 690
 Sequence: 1 ctgcacctgcagtgctga.....gctgaacccatctctat 690

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
 5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq.*
 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match %	Length	DB ID	Description
1	517.2	75.0	1225	4	US-09-016-434-1235
2	517.2	75.0	1317	1	US-08-153-848-45
3	517.2	75.0	1317	3	US-08-239-843A-45
4	517.2	75.0	1317	4	US-09-088-337B-45
5	517.2	75.0	1317	5	PCT-US93-11153-45
6	517.2	75.0	1664	4	US-09-582-224A-5
7	517.2	75.0	1664	4	US-09-023-655-1213
8	517.2	75.0	1679	4	US-09-517-805-14
9	517.2	75.0	1737	1	US-08-202-056-4
10	517.2	75.0	1737	1	US-08-076-093A-3
11	517.2	75.0	1737	1	US-08-701-265-3
12	517.2	75.0	1737	2	US-08-284-586-3
13	517.2	75.0	1737	2	US-08-805-478-3
14	517.2	75.0	1737	2	US-08-802-627A-3
15	517.2	75.0	1737	2	US-08-801-238-3
16	517.2	75.0	1737	2	US-08-801-228-3
17	517.2	75.0	1737	3	US-09-104-296-3
18	517.2	75.0	1737	5	PCT-US94-06380-2
19	150	21.7	1119	4	US-09-170-496D-65
20	150	21.7	1679	1	US-08-202-056-6
21	150	21.7	1679	1	US-08-076-093A-5
22	150	21.7	1679	1	US-08-701-265-5
23	150	21.7	1679	2	US-08-284-586-5
24	150	21.7	1679	2	US-08-805-478-5
25	150	21.7	1679	2	US-08-802-627A-5
26	150	21.7	1679	2	US-08-801-238-5
27	150	21.7	1679	2	US-08-801-228-5

ALIGNMENTS

RESULT 1

US-09-016-434-1235
 ; Sequence 1235, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1235:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1225 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g189313
 ; US-09-016-434-1235

Query Match 75.0%; Score 517.2; DB 4; Length 1225;
 Best Local Similarity 85.8%; Pred. No. 5.7e-140;

28	150	21.7	1679	3	US-09-104-296-5	Sequence 5, Appli
29	150	21.7	1679	5	PCT-US94-06380-3	Sequence 3, Appli
30	150	21.7	2818	3	US-08-982-493-7	Sequence 7, Appli
31	150	21.7	2818	3	US-08-628-655-1	Sequence 1, Appli
32	148.4	21.5	1119	4	US-09-170-496D-199	Sequence 199, App
33	142.2	20.6	1176	5	PCT-US95-03032-2	Sequence 2, Appli
34	142.2	20.6	1883	1	US-08-202-056-2	Sequence 2, Appli
35	142.2	20.6	1933	1	US-08-076-093A-1	Sequence 1, Appli
36	142.2	20.6	1933	1	US-08-410-451-1	Sequence 1, Appli
37	142.2	20.6	1933	1	US-08-410-455-1	Sequence 1, Appli
38	142.2	20.6	1933	1	US-08-418-919-1	Sequence 1, Appli
39	142.2	20.6	1933	1	US-08-410-453A-2	Sequence 2, Appli
40	142.2	20.6	1933	1	US-08-701-265-1	Sequence 2, Appli
41	142.2	20.6	1933	1	US-08-410-454A-2	Sequence 2, Appli
42	142.2	20.6	1933	2	US-08-284-586-1	Sequence 1, Appli
43	142.2	20.6	1933	2	US-08-410-456A-2	Sequence 2, Appli
44	142.2	20.6	1933	2	US-08-805-478-1	Sequence 1, Appli
45	142.2	20.6	1933	2	US-08-802-627A-1	Sequence 1, Appli

Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGCACTGTGAGTGGCTGACCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 60
Db 257 CTGCACTGTGAGTGGGAGCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 316
QY 61 GCCATGGGTGACTGTACTTTGGGAAATTTTGTAAAGGTGTCCATATCATCTACACT 120
Db 317 GCGGTGGCAAACTGGTACTTTGGGAACTTCTATSCAAGGAGTCCATGTCTACACA 376
QY 121 GTCAACTCTAGCAGAGGTTCATCTCTGGCTTCATCAGCTGGAGCGGTACTGGCC 180
Db 377 GTCAACTCTAGCAGAGGTTCATCTCTGGCTTCATCAGCTGGAGCGGTACTGGCC 436
QY 181 ATTGTCCAGCCCAACAGTCAAGAGGCAAGAACTGCTGGTGAAGGAGTCTAT 240
Db 437 ATCTGTCCAGCCCAACAGTCAAGAGGCAAGAACTGCTGGTGAAGGAGTCTAT 496
QY 241 GTGGGCTGTGAATCCAGCCCTCTCTCTGTGATATACCTGACTTCATCTTTGGCAGGTC 300
Db 497 GTTGGCTGTGAATCCAGCCCTCTCTCTGTGATATACCTGACTTCATCTTTGGCAGG-- 554
QY 301 AGCCAGGGGACATCAGTCAAGGAGGATGACAGTATCTGTGACCGCTTTTACCCGAT 360
Db 555 -----TCAGTGAGGACATGACAGATATCTGTGACCGCTTTTACCCGAT 601
QY 361 AGCCTGTGATGGTGGTGTTCATTTCCAGCATATATGGTGGGTCTCATCTGCCGGC 420
Db 602 GACTTGTGGTGGTGTGTTCAGTTTCAGCATATATGGTGGGTCTCATCTGCCGTGT 661
QY 421 ATCTGTATCTCTCTGTGATGATCATCTCTTAAAGTGTGACATCTCAAGGGCAC 480
Db 662 ATTGTATCTCTCTGTGATGATCATCTCTTAAAGTGTGACATCTCAAGGGCAC 721
QY 481 CAGAGGCAAGGCGCTCAAGAGCAGTATCTCTTCACTCTCTCTCTCTCTCTCTCT 540
Db 722 CAGAGGCAAGGCGCTCAAGAGCAGTATCTCTTCACTCTCTCTCTCTCTCTCTCT 781
QY 541 CTGCCATATATGTGGGATGAGTATGAGTATCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 782 CTGCTTACTATATGGGATGAGTATGAGTATCTCTCTCTCTCTCTCTCTCTCTCT 841
QY 601 GGATGTGACTTCGAGAGCTTGTGACAGTGGATCTCCATCAGAGGCGCTCTGCTTC 660
Db 842 GGGTGTGATTTGAGACACTGTGACAGTGGATTTCCATCAGGAGCGCTAGCTTTC 901
QY 661 TTCCTGTGTGCTGAACCCCTCTCTCTAT 690
Db 902 TTCCACTGTGTGCTGAACCCCTCTCTAT 931

RESULT 2

US-08-153-848-45
; Sequence 45, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; ADDRESS: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,848
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804 and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 201..1211
US-08-153-848-45
Query Match 75.0%; Score 517.2; DB 1; Length 1317;
Best Local Similarity 85.8%; Pred. No. 6e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGCACTGTGAGTGGCTGACCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 60
Db 387 CTGCACTGTGAGTGGGAGCTCTCTTTGTATCATCACTCCCTTCTGGGAGTTGAT 446
QY 61 GCCATGGGTGACTGTACTTTGGGAAATTTTGTAAAGGTGTCCATATCATCTACACT 120
Db 447 GCGGTGGCAAACTGGTACTTTGGGAACTTCTATSCAAGGAGTCCATGTCTACACA 506
QY 121 GTCAACTCTAGCAGAGGTTCATCTCTGGCTTCATCAGCTGGAGCGGTACTGGCC 180
Db 507 GTCAACTCTAGCAGAGGTTCATCTCTGGCTTCATCAGCTGGAGCGGTACTGGCC 566
QY 181 ATTGTCCAGCCCAACAGTCAAGAGGCAAGAACTGCTGGTGAAGGAGTCTAT 240
Db 567 ATCTGTATCTCTCTGTGATGATCATCTCTTAAAGTGTGACATCTCAAGGGCAC 626
QY 241 GTGGGCTGTGATCCAGCCCTCTCTCTGACTATACCTGACTTCATCTTTGGCAGCTC 300
Db 627 GTTGGCTGTGATCCCTGCTCTGCTGACTATTCCGACTTCATCTTTGCCAAG-- 684
QY 301 AGCCAGGGGACATCAGTCAAGGAGGATGACAGTATCTGTGACCGCTTTTACCCGAT 360
Db 685 -----TCAGTGAGGAGATGACAGATATCTGTGACCGCTTTTACCCGAT 731
QY 361 AGCCTGTGATGGTGGTGTTCATTTCCAGCATATATGGTGGGTCTCATCTGCCGGC 420
Db 732 GACTTGTGGTGGTGTGTTCAGATTTCAGCATATCATGTTGGCTTATCTCTGCTGGT 791
QY 421 ATCTGTATCTCTCTGTGATGATCATCTCTTAAAGTGTGACATCTCAAGGGCAC 480
Db 792 ATTGTATCTCTCTGTGATGATCATCTCTTAAAGTGTGACATCTCAAGGGCAC 851
QY 481 CAGAGGCAAGGCGCTCAAGAGCAGTATCTCTCTCATCTCTGCTTTCTTTGCTGCTG 540
Db 852 CAGAGGCAAGGCGCTCAAGAGCAGTATCTCTCTCATCTCTGCTTTCTTTGCTGCTG 911
QY 541 CTGCCATATATGTGGGATGAGTATGAGTATCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 912 CTGCTTACTATATGGGATGAGTATGAGTATCTCTCTCTCTCTCTCTCTCTCTCT 971
QY 601 GGATGTGACTTCGAGAGCTTGTGCAAGTGGATCTCTCATCAGAGGCGCTTCTGCTTC 660

Db 972 GGGTGTGAGTTTGGAGCACTGTGCACAGTGGATTTCATCCAGGCGCTTAGCTTTC 1031
Qy 661 TTCCACCTGTGGCTGACACCCCATCCTCTAT 690
Db 1032 TTCCACCTGTGGCTGACACCCCATCCTCTAT 1061

RESULT 3
US-09-299-843A-45
; Sequence 45, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,843A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/088,337
; FILING DATE: 01-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jill E. Uhl
; REGISTRATION NUMBER: 43,213
; REFERENCE/DOCKET NUMBER: 27866/32059B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 201..1211
; US-09-299-843A-45

Query Match 75.0%; Score 517.2; DB 3; Length 1317;
Best Local Similarity 85.8%; Pred. No. 6e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy 1 CTGACCTGTGAGTGGCTGACCTCTCTTTGTGTATCATCATCCCTTCTGGGAGTTGAT 60
Db 387 CTGACCTGTGAGTGGGCGACCTCTCTTTGTGTATCATCATCGTTCCCTTCTGGGAGTTGAT 446
Qy 61 GCCATGGCTGACTGTACTTTGGGAAATTTTGTGTAAGGCTGCCATCATCATCTACACT 120

Db 447 GCCGTGGCAAACTGGTACTTTTGGGAACTTCTCTATGCAAGGAGTCCATGTCTATCACA 506
Qy 121 GTCAACCTCTACAGCAGCGTTCTCATCTGCGCTTCTATCAGCTTGGACCGGTACTCGCC 180
Db 507 GTCAACCTCTACAGCAGTGTCTCATCTGCGCTTCTATCAGCTTGGACCGGTACTCGCC 566
Qy 181 ATTGTCCACGCCCAACACAGTCAAAGGCCAAAGGCAATCTGCTGAGTGAAGGCAAGTCTAT 240
Db 567 ATCGTCCACGCCCAACACAGTCAAAGGCCAAAGGCAATCTGCTGAGTGAAGGCAAGTCTAT 626
Qy 241 GTGGCGTCTGGATCCCGACCGCTCTCTGACTATACCTGACTTCTATCTTTTGGCGCGTC 300
Db 627 GTTGGCGTCTGGATCCCGACCGCTCTCTGACTATACCTGACTTCTATCTTTTGGCGCGTC 684
Qy 301 AGCCAGGCGGACATCAGTCAGGCGGATGACAGGATACATCTCTGACCGCTTTTACCCCGAT 360
Db 685 -----TCAGTGAGCGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 731
Qy 361 AGCTGTGGATGGTGGTGTCTTCAATTCAGCATATATATGTTGGTGTCTCATCTGCGCGC 420
Db 732 GACTTGTGGTGGTGTGTGTCTTCCAGTTTCAGCATATCATGTTGGTGTCTCATCTGCGCGC 791
Qy 421 ATCGTCATCTCTCTCTGTGTCTGCTATCTGCTATCTCTAAGCTGTGACATCTCCAAAGGCGCAC 480
Db 792 ATTGTATCTCTGTCTGTCTATCTGCTATCTCTAAGCTGTGACATCTCCAAAGGCGCAC 851
Qy 481 CAGAGCGCAAGGCGCTCTCAAGCAGCATATCTCTATCTCTGCTGTCTCTTTTGGCGCTGTC 540
Db 852 CAGAGCGCAAGGCGCTCTCAAGCAGCATATCTCTATCTCTGCTGTCTCTTTTGGCGCTGTC 911
Qy 541 CTGCCATATTATGTTGGGATCAGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 600
Db 912 CTGCCATATTATGTTGGGATCAGATCGATCGATCGATCGATCGATCGATCGATCGATCGAT 971
Qy 601 GGATGTGACTTTCGAGAGCATTTGTCACAAAGTGGATCTCCATCAGAGGCGCTCTGCGCTTC 660
Db 972 GGGTGTGAGTTTGGAGCACTGTGCAAGTGGATTTCCATCAGGCGCGCTTAGCTTTC 1031
Qy 661 TTCCACTGTGCTGCAACCCCATCTCTAT 690
Db 1032 TTCCACTGTGCTGCAACCCCATCTCTAT 1061

RESULT 4
US-09-088-337B-45
; Sequence 45, Application US/09088337B
; Patent No. 6348574
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6348574el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,337B
; FILING DATE: 01-JUN-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848

FILING DATE: 17-NOV-1993
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 6348574and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 201..1211
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-088-337B-45

Query Match 75.0%; Score 517.2; DB 4; Length 1317;
Best Local Similarity 85.8%; Pred. No. 6e-140; Indels 15; Gaps 1;
Matches 592; Conservative 0; Mismatches 83;

QY 1 CTGCACCTGTACAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCGAGTTGAT 60
DB 387 CTGCACCTGTACAGTGGCGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCGAGTTGAT 446
QY 61 GCATGGCTGACGTGGTACTTTGGGAAATTTTGTGTAGGCTGCCATATCATCTACACT 120
DB 447 GCGGTGGCAAACTGGTACTTTGGGAACTTCTTATGCAAGGCGAGTCCATGTCTACACA 506
QY 121 GTCAACCTCTACAGCAGCGTCTCTCTGCGCTTCATCAGCTGGACCGGTACCTCGCC 180
DB 507 GTCAACCTCTACAGCAGTGTCTCTCTGCGCTTCATCAGCTGGACCGGTACCTCGCC 566
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCGCAAGGAACTGTGCTGGCTGAAAGCGAGTCTAT 240
DB 567 ATCGTCCAGCCACCAACAGTCAAGGCGCAAGGAACTGTGCTGGCTGAAAGCGAGTCTAT 626
QY 241 GTGGGCGTCTGATCCAGCGCTCTCTCTGACTATACCTGACTTCTTTGCGGAGTCT 300
DB 627 GTGGGCGTCTGATCCCTGCGCTCTCTCTGACTATACCTGACTTCTTTGCGGAGTCT 684
QY 301 AGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTTACCCGAT 360
DB 685 -----TCAGTGAGGACATGACAGATATATCTGTGACCGCTTCTACCCCAAT 731
QY 361 AGCCTGTGATGGTGGTCTTCAATCCAGCATATAATGTTGGTCTCATCTGCGCGC 420
DB 732 GACTTGTGGTGGTGGTTCAGTTTCAGTTCAGACATCATGGTGGCTTATCTGCGTGT 791
QY 421 ATCGTCATCTCTCTGTTTACTGTCATCATCTCTTAAGCTGTGTCACACTCCAGGGCCAC 480
DB 792 ATTGTATCTCTCTGCTGTTGCTATTTATCTCTCAAGCTGTGTCACACTCCAGGGCCAC 851
QY 481 CAGAGCGGACGCGCTCAAGCAGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 852 CAGAGCGGACGCGCTCAAGCAGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 911
QY 541 CTGCCATATATATGTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 912 CTGCCATATATATGTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 971
QY 601 GGATGTGACTTCGAGAGCATTTGTGCAAGTGGATCTCTCATCAGAGGCGCTCGCTTC 660
DB 972 GGGTGTGATTTGAGAACATCTGTGCAAGTGGATTTCCATCAGAGGCGCTCGCTTC 1031
QY 661 TTCCACTGTGCTGCAACCCCATCTCTAT 690

DB 1032 TTCCACTGTGTCTGGAACCCCATCTCTAT 1061

RESULT 5
PCT-US93-11153-45
; Sequence 45, Application PC/TUS9311153
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schwellkart, Vicki L.
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11153
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 07/977,452
; APPLICATION NUMBER: 17-NOV-1992
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 201..1211
; PCT-US93-11153-45

Query Match 75.0%; Score 517.2; DB 5; Length 1317;
Best Local Similarity 85.8%; Pred. No. 6e-140; Indels 15; Gaps 1;
Matches 592; Conservative 0; Mismatches 83;

QY 1 CTGCACCTGTACAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCGAGTTGAT 60
DB 387 CTGCACCTGTACAGTGGCGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCGAGTTGAT 446
QY 61 GCATGGCTGACGTGGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACT 120
DB 447 GCGGTGGCAAACTGGTACTTTGGGAACTTCTTATGCAAGGCGAGTCCATGTCTACACA 506
QY 121 GTCAACCTCTACAGCAGCGTCTCTCTGCGCTTCATCAGCTGGACCGGTACCTCGCC 180
DB 507 GTCAACCTCTACAGCAGTGTCTCTCTGCGCTTCATCAGCTGGACCGGTACCTCGCC 566
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCGCAAGGAACTGTGCTGGCTGAAAGCGAGTCTAT 240
DB 567 ATCGTCCAGCCACCAACAGTCAAGGCGCAAGGAACTGTGCTGGCTGAAAGCGAGTCTAT 626

181	ATTGTCACGCCACCAACAGTCAAGGCCAAGGAACATGCTGGCTGGAGGAAGCGACTCTAT	240
488	ATCGTCCAGCCCAACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAGGGTGTCTAT	547
241	GTGGGGCTGTGGATCCACAGCCCTCCTCCGTGACTATACCTGACTTCATCTTTGCCGACGTC	300
548	GTTGGCGTCTGGATTCCCTGGCCCTCTGCTGACTATTCGCGACTTCATCTTTGCCAAG--	605
301	AGCCAGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGGCTTTTACCOCGAT	360
606	-----TCAGTGAGGCAGATGACAGATATCTGTGACGGCTTCTACGCCCAAT	652
361	AGCCTGTGGATGTGGTGTTTCAATTCACGACATATATGGTGGGTCTCATCTCGTCCCGGC	420
653	GACTGTGGGTGTGTGTTCAGTTTCAGACATCATGGTTGGGCTTATCTCGCTCTGT	712
421	ATCGTCATCCTCTCTGTACTGCATCATCATCTCTTAAGCTGTGCACATCCAAAGGGCCAC	480
713	ATTGTCACTCTGTCTCTGCTATTGCAATTATCATCTCCAAGCTGTCAACATCCAAAGGGCCAC	772
481	CAGAAGGCGAAGGCCCTCAAGACGAGTCATCTCATCTCCTTAGCTTCTTTGGCTGTGTGG	540
773	CAGAAGGCGAAGGCCCTCAAGACCAAGTCATCTCTCATCTCGGCTTCTTTCGCCCTGTTGG	832
541	CTGCGCATATTATGTGGGATCAGCATCGACTCTCTCATCTTTTGGGAGTCATCAAGCAA	600
833	CTGCTTACTACTATTGGATCAGCATCGACTCTTTCATCTCTCTCGGAATCATCAAGCAA	892
601	GGATGTGACTTCAGAGCAATTGTGCACAAGTGGATCTCCATCACAGAGGCCCTCGCCCTTC	660
893	GGGTGTGAGTTTCAGAAACACTGTGTCAACAAGTGGATTTCCATCACGAGAGCCCTAGCTTTC	952

953 TCCACTGTTGTCTGAACCCCATCTCTAT 982
 RESULT 7
 US-09-023-655-1213
 ; Sequence 1213, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:

REMOV. LIFE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/023,655
 FILING DATE: HEREWITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0001 US

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1213:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g219868
US-09-023-655-1213

Query Match
Best Local Similarity 75.0%; Score 517.2; DB 4; Length 1664;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

1 CTGCACCTGTGACGTGCTGACCTCCCTTTGTGATCAGACCTCCCTTCTGCGCAGTGTAT 60
308 CTGCACCTGTGACGTGCGGACCTCCCTTTGTGATCAGCCTCCCTTCTGCGCAGTGTAT 367
61 GCCATGCGCTGACGTGCTGACCTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
368 GCCGTGGCAAACTGGTACCTTTGGGAACTTCCATGCAAGGAGTCCATGTCTATCACA 427
121 GTCAACCTCTACAGACGCTTCTATCTGCGCTTCTATCAGCCTGAGCGGTACCTCGCC 180
428 GTCAACCTCTACAGACGCTGCTCTATCTGCGCTTCTATCAGCCTGAGCGGTACCTCGCC 487
181 ATTGTCCAGCCCAACAGCTCAAGCCCAAGGAACTGCTGCTGAAAGGAGCTCTAT 240
488 ATCGTCCAGCCCAACAGCTCAAGCCCAAGGAACTGCTGCTGAAAGGAGTGTCTAT 547
241 GTGGGCTGTGGATCCAGCGCTTCTCTGATATACCTGATCTTCTTTGCGGAGCTC 300
548 GTTGGGCTGTGGATCCAGCGCTTCTCTGATATACCTGATCTTCTTTGCGGAGCTC 605
301 AGCCAGGGGAGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
606 -----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 552
361 AGCCTGTGGATGTGGTGTTCATTTCCAGCATATATATGTTGGTCTCTATCTGCGCGGC 420
653 GACTTGTGGTGTGGTGTTCATTTCCAGCATATATGTTGGTCTCTATCTGCGCGGT 712
421 ATGCTATCTCTCTGTTACTGATCATCTCTAAGCTGTGACATCTCAAGGCGCAC 480
713 ATTGTATCTCTCTGTTACTGATCATCTCTAAGCTGTGACATCTCAAGGCGCAC 772
481 CAGAAGCGCAAGCGCTCAAGACGACAGTCACTCTATCTGCTGCTGCTGCTGCTG 540
773 CAGAAGCGCAAGCGCTCAAGACGACAGTCACTCTATCTGCTGCTGCTGCTGCTG 832
541 CTGCGCATATATGTTGGGATCAGATCGACTCTCTATCTGCTGCTGCTGCTGCTG 600
833 CTGCGCTTACTATGTTGGGATCAGATCGACTCTCTATCTGCTGCTGCTGCTGCTG 892
601 GGATGTGACTTCGAGAGCATTTGCAAGTGTGATCTCCATCAGAGGCGCTGCGCTTC 660
893 GGGTGTGAGTTGAGAACATCTGCAAGTGTGATTTCCATCAGAGGCGCTGAGCTTTC 952
661 TTCCACTGTGCTGAAACCCCATCTCTAT 690
953 TTCCACTGTGCTGAAACCCCATCTCTAT 982

RESULT 8
US-09-517-605-14
Sequence 14, Application US/09517605
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Littman, Dan R.
APPLICANT: Kwon, Douglas S.

APPLICANT: van Kooyk, Yvette
APPLICANT: Geijtenbeek, Theo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
CELLS
TITLE OF INVENTION: CELLS
FILE REFERENCE: 1049-1-017
CURRENT APPLICATION NUMBER: US/09/517,605
CURRENT FILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 1679
TYPE: DNA
ORGANISM: Homo sapiens
US-09-517-605-14

Query Match
Best Local Similarity 75.0%; Score 517.2; DB 4; Length 1679;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

1 CTGCACCTGTGACGTGCTGACCTCCCTTTGTGATCAGACCTCCCTTCTGCGCAGTGTAT 60
320 CTGCACCTGTGACGTGCGGACCTCCCTTTGTGATCAGCCTCCCTTCTGCGCAGTGTAT 379
61 GCCATGCGCTGACGTGCTGACCTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
380 GCCGTGGCAAACTGGTACCTTTGGGAACTTCCATGCAAGGAGTCCATGTCTATCACA 439
121 GTCAACCTCTACAGACGCTTCTATCTGCGCTTCTATCAGCCTGAGCGGTACCTCGCC 180
440 GTCAACCTCTACAGACGCTGCTCTATCTGCGCTTCTATCAGCCTGAGCGGTACCTCGCC 499
181 ATTGTCCAGCCCAACAGCTCAAGCCCAAGGAACTGCTGCTGAAAGGAGCTCTAT 240
500 ATCGTCCAGCCCAACAGCTCAAGCCCAAGGAACTGCTGCTGAAAGGAGTGTCTAT 559
241 GTGGGCTGTGGATCCAGCGCTTCTCTGATATACCTGATCTTCTTTGCGGAGCTC 300
560 GTTGGGCTGTGGATCCAGCGCTTCTCTGATATACCTGATCTTCTTTGCGGAGCTC 617
301 AGCCAGGGGAGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
618 -----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 664
361 AGCCTGTGGATGTGGTGTTCATTTCCAGCATATATATGTTGGTCTCTATCTGCGCGGC 420
665 GACTTGTGGTGTGGTGTTCATTTCCAGCATATATGTTGGTCTCTATCTGCGCGGT 724
421 ATGCTATCTCTCTGTTACTGATCATCTCTAAGCTGTGACATCTCAAGGCGCAC 480
725 ATTGTATCTCTCTGTTACTGATCATCTCTAAGCTGTGACATCTCAAGGCGCAC 784
481 CAGAAGCGCAAGCGCTCAAGACGACAGTCACTCTATCTGCTGCTGCTGCTGCTG 540
785 CAGAAGCGCAAGCGCTCAAGACGACAGTCACTCTATCTGCTGCTGCTGCTGCTG 844
541 CTGCGCATATATGTTGGGATCAGATCGACTCTCTATCTGCTGCTGCTGCTGCTG 600
845 CTGCGCTTACTATGTTGGGATCAGATCGACTCTCTATCTGCTGCTGCTGCTGCTG 904
601 GGATGTGACTTCGAGAGCATTTGCAAGTGTGATCTCCATCAGAGGCGCTGCGCTTC 660
905 GGGTGTGAGTTGAGAACATCTGCAAGTGTGATTTCCATCAGAGGCGCTGAGCTTTC 964
661 TTCCACTGTGCTGAAACCCCATCTCTAT 690
965 TTCCACTGTGCTGAAACCCCATCTCTAT 994

RESULT 9
US-08-202-056-4
Sequence 4, Application US/08202056
Patent No. 5440021
GENERAL INFORMATION:
APPLICANT: Kwon, Douglas S.

APPLICANT: Chuntharapai, Anan
APPLICANT: Hebert, Caroline
APPLICANT: Kim, Kyung Jin
APPLICANT: Lee, James
TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,056

FILING DATE: 25-FEB-1994
CLASSIFICATION: 436

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211

FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:

NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659

REFERENCE/DOCKET NUMBER: 706P3
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881

TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 1737 bases

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
US-08-202-056-4

Query Match 75.0%; Score 517.2; DB 1; Length 1737;
Best Local Similarity 85.8%; Pred. No. 6.8e-140;

Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

Qy 1 CTGCACCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 60
Db 322 CTGCACCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 381
Qy 61 GCCATGCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 120
Db 382 GCCATGCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 441
Qy 121 GTCAACCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 180
Db 442 GTCAACCTGTACAGCGTGTCTTCTTGTGTCATCACACTCCCTCTTGGGCGAGTTGAT 501
Qy 181 ATTGTCCACGCCACCAACAGCTCAAGAGCCCAAGAAAGCTGTGCTGCTGCTGCTGCTGCT 240
Db 502 ATGCTCAGCGCCACCAACAGCTCAAGAGCCCAAGAAAGCTGTGCTGCTGCTGCTGCTGCT 561
Qy 241 GTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 562 GTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
Qy 301 AGCCAGGGGACATCAGTCAGGGGGGATGACAGGTATCATCTGTACCGCCCTTTACCCCGAT 360
Db 620 -----TCAGTGAGGAGATGACAGGTATCATCTGTACCGCCCTTTACCCCGAT 666
Qy 361 AGCTGTGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 420
Db 667 GACTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 726

Qy 421 ATCGTCATCCTCTCTCTGTTACTGTCATCATCTCTTAAGTGTGTCACATCTCAAGGGCCAC 480
Db 727 ATTGTGTCATCCTCTCTGTTACTGTCATCATCTCTTAAGTGTGTCACATCTCAAGGGCCAC 786
Qy 481 CAGAAGCGCAAGGGCCCTCAAGACGACAGTCTCTCTCATCTCTTCTTCTTCTTCTTCTTCTTCT 540
Db 787 CAGAAGCGCAAGGGCCCTCAAGACGACAGTCTCTCTCATCTCTTCTTCTTCTTCTTCTTCTTCT 846
Qy 541 CTGCCATATTATGTTGGGATCAGATCAGATCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
Db 847 CTGCCATATTATGTTGGGATCAGATCAGATCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 906
Qy 601 GGATGTGACTTCGAGAGCATTTGTGCAAGTGGATCTCCATCAGAGAGGCGCTTCGCCCTTC 660
Db 907 GGGTGTGAGTTTGAGAACACTGTGCAAGTGGATTTCCATCAGAGAGGCGCTTCGCCCTTC 966
Qy 661 TTCCACTGTTGCTGAGAACCCCATCTCTTAT 690
Db 967 TTCCACTGTTGCTGAGAACCCCATCTCTTAT 996

RESULT 10

US-08-076-093A-3
Sequence 3, Application US/08076093A

Patent No. 5543503
GENERAL INFORMATION:

APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James

APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.

TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco

STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A

FILING DATE: 11-Jun-1993
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782

FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.

REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530

TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1737 nucleotides
TYPE: Nucleic Acid

STRANDEDNESS: Single
TOPOLOGY: Linear

US-08-076-093A-3

Query Match 75.0%; Score 517.2; DB 1; Length 1737;
Best Local Similarity 85.8%; Pred. No. 6.8e-140;

Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY	1	CTGCACCTGTGACGTGGCTGACCTCCTCTTTGTGTAACACATCTCCCTTCTGGGACGTTGAT	60
Db	322	CTGCACCTGTGACGTGGCGACCTCCTCTTTGTGTAACACGCTTCCCTTCTGGGACGTTGAT	381
QY	61	GCCATGGCTGACTGGTGACTTTTGGGAAATTTTGTGTAAAGGCTGTCCATATCATCTACACT	120
Db	382	GCGGTGGCAAACTGCTACTTTTGGGAACTTCTCTATGACGAGCACTCCATGTCATCTACACA	441
QY	121	GTCAACCTCTACAGCAGCGTTCTCATCTGGCCTTCACTCAGCTGACCGGTACCTCGCC	180
Db	442	GTCAACCTCTACAGCAGGTGCTCATCTGGCCTTCACTCAGCTGACCGGTACCTCGCC	501
QY	181	ATTGTCCAGCCACCAACAGTCAAAAGGCCAAGGAACTGCTGGCTGAAAGGCAGTCTAT	240
Db	502	ATCGTCCAGCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAGGTGGTCTAT	561
QY	241	GTGGGGTCTGATCCAGCCCTCCTCCTGACTATACCTGACTTCATCTTTGGCAGCGTC	300
Db	562	GTGGGGTCTGATCCCGTCCCTCCTGCTGACTATCCGACTTCATCTTTGGCAACG--	619
QY	301	AGCCAGGGGACATCAGTCAGGGGATCACAGGTACATCTGTGACCGCCTTTACCCCGAT	360
Db	620	-----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT	666
QY	361	AGCCTGTGATGGTGGTGTTCATNTCAGACATATAATGTGGGTCTCATCTGCGCCGC	420
Db	667	GACTTGTGGTGGTGTGTGTGTCAGTTTCAGACATCATGGTGGCCCTATCTGCGCTGT	726
QY	421	ATCGTCATCCTCTCCTGTTACTGTCATCATCTCTAAGCTTCACATCCCAAGGGCCAC	480
Db	727	ATTGTATCCTGTCTCTGCTATTGCAATTATCATCTCCAAGTGTACACTCCAAGGCCAC	786
QY	481	CAGAAGCGAAGCCCTCAGACGACGTCTATCCTCATCTAGCTTTCTTTGGCTCGTG	540
Db	787	CAGAAGCGAAGCCCTCAAGACCACAGTCACTCCTCATCTGCGCTTTCTTTCGCGCTGTGG	846
QY	541	CTGCCATATTATGTGGGATCAGCATCGACTCCTTTCATCTTTTGGAGTCATCAAGCAA	600
Db	847	CTGCCTTACTATTGGGATCAGCATCGACTCCTTTCATCTCTCTGGAAATCATCAAGCAA	906
QY	601	GGATGTGACTTCGAGAGCATGTGCAGAAGTGGATCTCCATCAGAGGCCCTCGCCTTC	660
Db	907	GGGTGTGAGTTTGAGAACACTGTGCAGAAGTGGATTTCCATCACCGAGGCCCTAGCTTTC	966
QY	661	TTCCACTGTTGGCTGAACCCCATCCTCTAT	690
Db	967	TTCCATGTTGTTCTGAAGCCCATCCTCTAT	996

```

RESULT 11
US-08-701-265-3
; Sequence 3, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Km, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)

```


Db 907 GGGTGTGAGTTTGAGAACACTGTGCACAGTGGATTTCATCAGCGGCCCTAGCTTTC 966
QY 661 TTCCACTGTGCTGAAACCCCATCTCTAT 690
Db 967 TTCCACTGTGCTGAAACCCCATCTCTAT 996

RESULT 12
US-08-284-586-3
; Sequence 3, Application US/08284586
; Patent No. 5840856
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,586
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,093A
; FILING DATE: 11-Jun-1993
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-284-586-3

Query Match 75.0%; Score 517.2; DB 2; Length 1737;
Best Local Similarity 85.8%; Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGCACCTGTGAGTGGTGCACCTCTCTTTGTATCACAACCTCCCTTCTGGGCGATTGAT 60
Db 322 CTGCACCTGTGAGTGGGCGACCTCTCTTTGTATCACAACCTCTCTCTCTGGGCGATTGAT 391

QY 61 GCCATGCTGACTGTGACTTTGGGAAATTTTGTGTAAGCTGTCCATCATCTACACT 120
Db 382 GCCGTGGCAAACTGGTACTTTGGGAACCTCTCTATGCAAGCGAGTCCATGTCTACACA 441

QY 121 GTCAACCTCTACACAGCGTCTCTATCTGCGCTTCTATCAGCTGGACGCGTACCTCGCC 180

Db 442 GTCAACCTCTACAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 501
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAACTCTGCTGCTGAAAGCGAGTCTAT 240
Db 502 ATCTGTCACGCCACCAACAGTCAAGGCCAAGGAACTCTGCTGCTGAAAGCGAGTCTAT 561
QY 241 GTGGCGTCTGGATCCAGCCCT 300
Db 562 GTTGGCGTCTGGATCCCTGCT 619
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTCACTCTGTCAGTCTCTCTCTCTCTCTCT 360
Db 620 -----TCAGTGAGGCAGATGACATATCTGTGACCGCTTCTACCCCAAT 666
QY 361 AGCTGTGGATGGTGGTGTTCATATTCAGCATATAAATGCTGGTCTCTCTCTCTCTCTCT 420
Db 667 GACTTGTGGTGGTGTGTTCAGTTCAGCATATGCTGTGGTCTCTCTCTCTCTCTCTCTCT 726
QY 421 ATCGTCATCT 480
Db 727 ATTGTATCT 786
QY 481 CAGAAGCGCAAGGCCCTCAAGACGACAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 787 CAGAAGCGCAAGGCCCTCAAGACGACAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 846
QY 541 CTGCCATATTATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 847 CTGCCCTTACTACATTTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
QY 601 GGATGTGACTTCGAGAGCAATGTGCACAGTGGATCTCCATCAGAGGCGCCCTCGCCTTC 660
Db 907 GGGTGTGAGTTTGAGAACACTGTGCACAGTGGATTTCATCACCAGGCGCCCTAGCTTTC 966

661 TTCCACTGTGCTGAAACCCCATCTCTAT 690
967 TTCCACTGTGCTGAAACCCCATCTCTAT 996

RESULT 13
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,478
; FILING DATE: 25-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:

```
/ APPLICATION NUMBER: 07/810782
/ FILING DATE: 19-DEC-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Love, Richard B.
/ REGISTRATION NUMBER: 34,659
/ REFERENCE/DOCKET NUMBER: P07062P1C1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-5530
/ TELEFAX: 415/952-9881
/ TELEX: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1737 base pairs
/ TYPE: Nucleic Acid
/ STRANDEDNESS: Single
/ TOPOLOGY: Linear
/ US-08-805-478-3

Query Match          75.0%; Score 517.2; DB 2; Length 1737;
Best Local Similarity 85.8%; Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTCGACCTCTCAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCAGTTGAT 60
Db 322 CTCGACCTCTCAGTGGCGCCACCTCTCTTTGTGTCATCACAGCTTCCTTCTGGGCAGTTGAT 381
QY 61 GCATGGCTCAGTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACT 120
Db 382 GCGGTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGCAGTCCATGTCATCTACACA 441
QY 121 GTCAACCTCTACAGCAGCTTCTCTCCCTGGCTTCTATCAGCTGGAACGCTGACCTCGCC 180
Db 442 GTCAACCTCTACAGCAGTGTCTCTCTCTGCTTCTATCAGCTGTGGACCTACCTGGCC 501
QY 181 ATTGTCCAGCCCAACCAAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTCTAT 240
Db 502 ATCGTCCAGCCCAACCAAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTGTCTAT 561
QY 241 GTGGGGTCTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 562 GTGGGGTCTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 619
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCGAT 360
Db 620 -----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 666
QY 361 AGCTGTGATGCTGGTGTTCATTCAGCATATAATGTTGGTCTCATCTGCTCCCGGC 420
Db 667 GACTGTGGGTGTGTGTTCCAGTTTACGACATCATGTTGGCTTATCTCTGCTGTGT 726
QY 421 ATCGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 727 ATTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 786
QY 481 CAGAGCGCAGGCGCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 787 CAGAGCGCAGGCGCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 846
QY 541 CTGCCATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 847 CTGCCATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
QY 601 GATGTGATCTCGAGAGCATTGTGACAGAGTGGATCTCGATCAGAGGCGCTGCGCTTC 660
Db 907 GGGTGTGAGTTTGAGAACACTGTGACAGAGTGGATTTCCATCAGCGAGCGCTAGCTTTC 966
QY 661 TTCCACTGTGTGCTGAACCCCATCTCTAT 690
Db 967 TTCCACTGTGTGCTGAACCCCATCTCTAT 996
```

RESULT 14

US-08-802-627A-3

```
/ Sequence 3, Application US/08802627A
/ Patent No. 5892017
/ GENERAL INFORMATION:
/ APPLICANT: Lee, James
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/802,627A
/ FILING DATE: 19-Feb-1997
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/284586
/ FILING DATE: 10-AUG-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/076093
/ FILING DATE: 11-JUN-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/810782
/ FILING DATE: 19-DEC-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Love, Richard B.
/ REGISTRATION NUMBER: 34,659
/ REFERENCE/DOCKET NUMBER: P07062P1D2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-5530
/ TELEFAX: 415/952-9881
/ TELEX: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1737 base pairs
/ TYPE: Nucleic Acid
/ STRANDEDNESS: Single
/ TOPOLOGY: Linear
/ US-08-802-627A-3

Query Match          75.0%; Score 517.2; DB 2; Length 1737;
Best Local Similarity 85.8%; Pred. No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTCGACCTCTCAGTGGCTGACCTCTCTTTGTGTCATCACACTCCCTTCTGGGCAGTTGAT 60
Db 322 CTCGACCTCTCAGTGGCGCCACCTCTCTTTGTGTCATCACAGCTTCCTTCTGGGCAGTTGAT 381
QY 61 GCATGGCTCAGTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACT 120
Db 382 GCGGTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGCAGTCCATGTCATCTACACA 441
QY 121 GTCAACCTCTACAGCAGCTTCTCTCCCTGGCTTCTATCAGCTGGAACGCTGACCTCGCC 180
Db 442 GTCAACCTCTACAGCAGTGTCTCTCTCTGCTTCTATCAGCTGTGGACCTACCTGGCC 501
QY 181 ATTGTCCAGCCCAACCAAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTCTAT 240
Db 502 ATCGTCCAGCCCAACCAAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTGTCTAT 561
QY 241 GTGGGGTCTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 562 GTGGGGTCTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 619
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCGAT 360
Db 620 -----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 666
QY 361 AGCTGTGATGCTGGTGTTCATTCAGCATATAATGTTGGTCTCATCTGCTCCCGGC 420
Db 667 GACTGTGGGTGTGTGTTCCAGTTTACGACATCATGTTGGCTTATCTCTGCTGTGT 726
QY 421 ATCGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 727 ATTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 786
QY 481 CAGAGCGCAGGCGCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 787 CAGAGCGCAGGCGCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 846
QY 541 CTGCCATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 847 CTGCCATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 906
QY 601 GATGTGATCTCGAGAGCATTGTGACAGAGTGGATCTCGATCAGAGGCGCTGCGCTTC 660
Db 907 GGGTGTGAGTTTGAGAACACTGTGACAGAGTGGATTTCCATCAGCGAGCGCTAGCTTTC 966
QY 661 TTCCACTGTGTGCTGAACCCCATCTCTAT 690
Db 967 TTCCACTGTGTGCTGAACCCCATCTCTAT 996
```

```

Db 620 -----TCAGTGGGAGAGTACAGATATCTGTGACCGCTTCTACCCCAAT 666
QY 361 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATAATGTGGGTCTCTACCTGCGCGG 420
Db 667 GACTGTGGGTGGTGTGTTTCCAGTTCAGCACATCATGGTGGCCCTTATCTGCTGGT 726
QY 421 ATCTGATCTCTCTCTGTTACTGATCATCTTAAGCTGTGACACTCAAGGGCCAC 480
Db 727 ATTGTCTATCTCTCTGCTATGTAATATCATCTCCAAAGTGTGACACTCCAAAGGGCCAC 786
QY 481 CAGAAGGCGAAGGCGCTCAAGACAGACAGTCTATCTCTATCTCTGCTGCTGCTGG 540
Db 787 CAGAAGGCGAAGGCGCTCAAGACAGACAGTCTATCTCTATCTCTGCTGCTGCTGG 846
QY 541 CTGCCATATATGTGGGATCAGATGATCTCTCTCTATCTCTGCTGCTGCTGG 600
Db 847 CTGCCCTTACTACATATGGGATCAGATGATCTCTCTCTATCTCTGCTGCTGCTGG 906
QY 601 GGATGTGACTTCGAGAGCATTTGACACAGTGGATCTCCATCAGAGGCGCTGCGCTTC 660
Db 907 GGGTGTGATTTGAGACACTGTGACACAGTGGATTTCCATCAGGAGGCGCTGCTTC 966
QY 661 TTCACACTGTTGCTGAACCCCATCTCTAT 690
Db 967 TTCACACTGTTGCTGAACCCCATCTCTAT 996

```

RESULT 15

```

US-08-801-238-3
; Sequence 3, Application US/08801238
; Patent No. 5919896
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PFAA RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.238
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs

```

```

; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-801-238-3
Query Match 75.0%; Score 517.2; DB 2; Length 1737;
Best Local Similarity 85.8%; Pred No. 6.8e-140;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGACCTGTGAGTGGGTGACCTCTCTCTTTGTGATCAGCACTCCCTCTTGGGCACTTGAAT 60
Db 322 CTGACCTGTGAGTGGGTGACCTCTCTCTTTGTGATCAGCACTCCCTCTTGGGCACTTGAAT 381
QY 61 GCATGGCTGACTGTGCTGCTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
Db 382 GCCTGGGCAACTGTGCTGCTTTGGGAACTTCTATGCAAGGAGTCCATGTCTATCTACACA 441
QY 121 GTCAACCTCTACAGCAGGCTTCTATCTCTGCTCTTATCAGCCTCGAAGCGTACCTGCGC 180
Db 442 GTCAACCTCTACAGCAGGCTTCTATCTCTGCTCTTATCAGTCTGAGCCGCTACCTGGCC 501
QY 181 ATGTGTCAGCGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAAAGGAGTCTAT 240
Db 502 ATGTCAGCGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAAAGGAGTCTAT 561
QY 241 GTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCTATCTTTGCCGACGTC 300
Db 562 GTTGGGCTCTGGATCCCTGCTCTCTCTGCTGACTATTTCCGACTTCTATCTTTGCCAAG-- 619
QY 301 AGCCAGGGGACATCAGTCAAGGGGATCAGAGTACATCTGTGACCGGCTTTACCCCGAT 360
Db 620 -----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 666
QY 361 AGCCTGTGGATGGTGTGTTTCAATTTCCAGCATATAATGGTGGGTCTCATCTCTGCCCGC 420
Db 667 GACTGTGGGTGGTGTGTTTCCAGTTCAGCACATCATGGTGGCCCTTATCTGCTGGT 726
QY 421 ATGTCATCTCTCTCTGTTACTGATCATCATCTCTTAAGCTGTGACACTCCAAAGGGCCAC 480
Db 727 ATTGTCTCTCTCTCTGCTATTTGCTATCTCTCAAGTGTGACACTCCAAAGGGCCAC 786
QY 481 CAGAAGGCGAAGGCGCTCAAGACAGACAGTCTCTCTCATCTCTGCTGCTGCTGG 540
Db 787 CAGAAGGCGAAGGCGCTCAAGACAGACAGTCTCTCTCATCTCTGCTGCTGCTGG 846
QY 541 CTGCCATATATGTGGGATCAGCATCGACTCTCTCTCTCTCTTTGGGAGTCTATCAAGCAA 600
Db 847 CTGCCCTTACTACATTTGGGATCAGCATCGACTCTCTCTCTCTCTGAAATCATCAAGCAA 906
QY 601 GGATGTGACTTCGAGAGCATTTGACACAGTGGATCTCCATCAGAGGCGCTTCCGCTTC 660
Db 907 GGGTGTGATTTGAGAACACTGTGACACAGTGGATTTCCATCAGGAGGCGCTTCTTC 966
QY 661 TTCACACTGTTGCTGAACCCCATCTCTAT 690
Db 967 TTCACACTGTTGCTGAACCCCATCTCTAT 996

```

Search completed: May 17, 2004, 16:15:43
Job time : 59.2446 secs

Sequence 14, Appl
Sequence 912, App
Sequence 332, App
Sequence 75, Appl
Sequence 58, Appl
Sequence 177, App
Sequence 20, Appl
Sequence 459, App
Sequence 3, Appl
Sequence 143, App
Sequence 143, App
Sequence 484, App
Sequence 674, App
Sequence 789, App
Sequence 687, App
Sequence 2611, Ap
Sequence 2611, Ap
Sequence 2611, Ap
Sequence 1, Appl
Sequence 1, Appl
Sequence 14955, A
Sequence 14955, A
Sequence 45, Appl
Sequence 182, App
Sequence 65, Appl
Sequence 5, Appl
Sequence 161, App
Sequence 59, Appl

15 517.2 75.0 1679 13 US-10-151-274-14
16 517.2 75.0 1679 13 US-10-172-118-912
17 517.2 75.0 1679 13 US-10-170-385-332
18 517.2 75.0 1679 13 US-10-225-567A-75
19 517.2 75.0 1679 15 US-10-021-860-58
20 517.2 75.0 1679 16 US-10-341-434-177
21 517.2 75.0 1711 10 US-09-971-392-20
22 517.2 75.0 1711 15 US-10-101-510-459
23 517.2 75.0 1737 9 US-09-104-063-3
24 517.2 75.0 1737 13 US-10-666-689-3
25 517.2 75.0 1737 13 US-09-870-759-143
26 517.2 75.0 1737 13 US-09-870-759-143
27 517.2 75.0 1737 13 US-09-870-759-143
28 517.2 75.0 1737 13 US-09-870-759-143
29 517.2 75.0 1737 13 US-09-870-759-143
30 517.2 75.0 1737 13 US-09-870-759-143
31 517.2 75.0 1737 13 US-09-870-759-143
32 517.2 75.0 1737 13 US-09-870-759-143
33 517.2 75.0 1737 13 US-09-870-759-143
34 517.2 75.0 1737 13 US-09-870-759-143
35 517.2 75.0 1737 13 US-09-870-759-143
36 517.2 75.0 1737 13 US-09-870-759-143
37 517.2 75.0 1737 13 US-09-870-759-143
38 517.2 75.0 1737 13 US-09-870-759-143
39 517.2 75.0 1737 13 US-09-870-759-143
40 517.2 75.0 1737 13 US-09-870-759-143
41 517.2 75.0 1737 13 US-09-870-759-143
42 517.2 75.0 1737 13 US-09-870-759-143
43 517.2 75.0 1737 13 US-09-870-759-143
44 517.2 75.0 1737 13 US-09-870-759-143
45 517.2 75.0 1737 13 US-09-870-759-143

OM nucleic - nucleic search, using sw model
Run on: May 17, 2004, 12:11:00 ; Search time 337.746 Seconds
(without alignments)
9271.041 Million cell updates/sec
Title: US-09-367-052-3
Perfect score: 690
Sequence: 1 ctgcaactgtcagtgctga.....gcctgaacccatctcttat 690
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2947324 seqs, 2269024515 residues
Total number of hits satisfying chosen parameters: 5894648
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
6: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
18: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
19: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *

ALIGNMENTS

RESULT 1
US-10-160-401-1
; Sequence 1, Application US/10160401
; Publication No. US20030207281A1
; GENERAL INFORMATION:
; APPLICANT: Genentech Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bieglecki, Karyn M.
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; APPLICANT: Rounds, Eileen
; TITLE OF INVENTION: HAPLOTYPES OF THE CXCR4 GENE
; FILE REFERENCE: MMH-0121US
; CURRENT APPLICATION NUMBER: US/10160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3118)..(3118)
; OTHER INFORMATION: PS1: polymorphic base thymine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3583)..(3583)
; OTHER INFORMATION: PS2: polymorphic base cytosine or thymine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3952)..(3952)
; OTHER INFORMATION: PS3: polymorphic base cytosine or thymine

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	75.1	5161	16	US-10-160-401-1
2	517.2	75.0	1059	13	US-09-813-651B-84
3	517.2	75.0	1059	13	US-10-014-322A-125
4	517.2	75.0	1059	16	US-10-160-401-2
5	517.2	75.0	1225	15	US-10-101-510-741
6	517.2	75.0	1225	15	US-10-305-720-1235
7	517.2	75.0	1664	17	US-10-641-643-1213
8	517.2	75.0	1670	9	US-09-880-107-2143
9	517.2	75.0	1670	10	US-09-960-706-636
10	517.2	75.0	1670	16	US-10-372-683-1
11	517.2	75.0	1670	16	US-10-440-464-77
12	517.2	75.0	1679	13	US-10-211-462-80
13	517.2	75.0	1679	13	US-10-181-906-9
14	517.2	75.0	1679	13	US-10-342-887-912

RESULT 2
US-09-813-651B-84
; Sequence 84, Application US/09813651B
; Publication NO. US20030018438A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Hehir, Christina
; APPLICANT: Kates, Steven
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds

RESULT 2
US-09-813-651B-84
; Sequence 84, Application US/09813651B
; Publication NO. US20030018438A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Hehir, Christina
; APPLICANT: Kates, Steven
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds

697 CAGAACGCAAGGCGCTCAAGACACACAGTCATCTCTGCTTTCTTGGCTTTGG 756
QY 541 CTGCCATATTATGTTGGGATCAGATCGACTCTCTTCACTCTTTGGGAGTATCAAGCAA 600
Db 757 CTGCCCTTACTATGTTGGGATCAGATCGACTCTCTTCACTCTTTGGGAGTATCAAGCAA 816
QY 601 GGATCTGACTTGGAGAGCATTTGTGCACAAAGTGGATCTCCATCAGAGAGGCTTCCCTTC 660
Db 817 GGGTGTGAGTTTGAGAACACTGTGCACAAAGTGGATTTCCATCAGAGAGGCTTCCCTTC 876
QY 661 TTCCACTCTGCTGAAACCCCATCTCTAT 690
Db 877 TTCCACTCTGCTGAAACCCCATCTCTAT 906

RESULT 4
US-10-160-401-2
; Sequence 2, Application US/10160401
; Publication No. US20030207281A1
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bieglecki, Karyn M.
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; APPLICANT: Rounds, Eileen
; TITLE OF INVENTION: HAPLOYPES OF THE CXCR4 GENE
; FILE REFERENCE: MMH-0121US
; CURRENT APPLICATION NUMBER: US/10/160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-160-401-2

Query Match 75.0%; Score 517.2; DB 16; Length 1059;
Best Local Similarity 85.8%; Pred. No. 1.2e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGCACTGTCTGAGTGGCTGACCTCTCTTTGTATCATCACACTCCCTCTTGGGAGTAT 60
Db 232 CTGCACTGTCTGAGTGGCGGACCTCTCTTTGTATCATCACACTCTCTTGGGAGTAT 291
QY 61 GCATGGCTGACTGTGACTTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
Db 292 GCCTGGCAAACTGTGACTTTTGGGAACTTCTATGCAAGGAGTCCATGTATCTACACA 351
QY 121 GTCAACCTCTACAGAGGTTCTCATCTCTGGCTTTCATCAGCTGGAGCGGTACCTCGCC 180
Db 352 GTCAACCTCTACAGAGTGTCTCATCTCTGGCTTTCATCAGCTGGAGCGGTACCTCGCC 411
QY 181 ATTGTCCAGCCCAACAGCTCAAGGCAAGGAACTGTGGCTGAAAGGAGTCTAT 240
Db 412 ATCTGCCAGCCCAACAGCTCAAGGCAAGGAACTGTGGCTGAAAGGAGTCTAT 471
QY 241 GTGGCGTCTGGATCCAGCCCTCTCTCTGACTATATCTGACTTCTATCTTTGGCGAGTC 300
Db 472 GTTGGCGTCTGGATCCCTGCTCTCTCTGACTATATCTGACTTCTATCTTTGGCGAGTC 529
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTATCTGTGCGGCTTACCCCGAT 360
Db 530 -----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 576
QY 361 AGCTGTGGATGGTGGTGTTCATATTCAGGATATAATGGTGGGTCTCATCTTGGCGGCG 420
Db 577 GACTTGTGGTGGTGTGTTCAGTTCAGGATATGAGTGGGCTTATCTCTGCTGTGT 636

US-10-014-322A-125
; Sequence 125, Application US/10014322A
; Publication No. US20030167129A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, Jr., John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Hehir, Christina
; APPLICANT: Kates, Steven
; APPLICANT: Krestenansky, John
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds
; FILE REFERENCE: CNS-008
; CURRENT APPLICATION NUMBER: US/10/014,322A
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/243,587
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 09/813,651
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 09/813,653
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 09/813,448
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(1059)
; OTHER INFORMATION: human CXCR4 gene
US-10-014-322A-125

Query Match 75.0%; Score 517.2; DB 15; Length 1059;
Best Local Similarity 85.8%; Pred. No. 1.2e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGCACTGTCTGAGTGGCTGACCTCTCTTTGTATCATCACACTCCCTCTTGGGAGTAT 60
Db 232 CTGCACTGTCTGAGTGGCGGACCTCTCTTTGTATCATCACACTCTCTTGGGAGTAT 291
QY 61 GCATGGCTGACTGTGACTTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
Db 292 GCCTGGCAAACTGTGACTTTTGGGAACTTCTATGCAAGGAGTCCATGTATCTACACA 351
QY 121 GTCAACCTCTACAGAGGTTCTCATCTCTGGCTTTCATCAGCTGGAGCGGTACCTCGCC 180
Db 352 GTCAACCTCTACAGAGTGTCTCATCTCTGGCTTTCATCAGCTGGAGCGGTACCTCGCC 411
QY 181 ATTGTCCAGCCCAACAGCTCAAGGCAAGGAACTGTGGCTGAAAGGAGTCTAT 240
Db 412 ATCTGCCAGCCCAACAGCTCAAGGCAAGGAACTGTGGCTGAAAGGAGTCTAT 471
QY 241 GTGGCGTCTGGATCCAGCCCTCTCTCTGACTATATCTGACTTCTATCTTTGGCGAGTC 300
Db 472 GTTGGCGTCTGGATCCCTGCTCTCTCTGACTATATCTGACTTCTATCTTTGGCGAGTC 529
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTATCTGTGCGGCTTACCCCGAT 360
Db 530 -----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 576
QY 361 AGCTGTGGATGGTGGTGTTCATATTCAGGATATAATGGTGGGTCTCATCTTGGCGGCG 420
Db 577 GACTTGTGGTGGTGTGTTCAGTTCAGGATATGAGTGGGCTTATCTTGGCGGCTGT 636
QY 421 ATCTGATCTCTCTGTTTATGATCATCATCTCTAAGTGTGCACACTCAAGGGCCAC 480
Db 637 ATTGTATCTCTCTGTTTATGATCATCTCTAAGTGTGCACACTCAAGGGCCAC 696
QY 481 CAGAGCCCAAGGCGCTCAAGGAGGAGTCACTCTCTCTGAGTCTTCTTGGCTGTGT 540

QY 421 ATCGTCATCTCTCTGTTTACTGATCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 480
DB 637 ATTGTTCATCTCTCTGTTTACTGATCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 696
QY 481 CAGAAGCGCAAGCCCTCAAGACGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 540
DB 697 CAGAAGCGCAAGCCCTCAAGACGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 756
QY 541 CTGCCATATTATGTGGGATCAAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 600
DB 757 CTGCCATATTATGTGGGATCAAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 816
QY 601 GGATGTGATCTTCAGAGCATTTGGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 660
DB 817 GGATGTGATCTTCAGAGCATTTGGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 876
QY 661 TTCCACTGTTGCTGAACCCCATCTCTAT 690
DB 877 TTCCACTGTTGCTGAACCCCATCTCTAT 906

RESULT 5
US-10-101-510-741
; Sequence 741, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 741
; LENGTH: 1225
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-741

Query Match 75.0%; Score 517.2; DB 15; Length 1225;
Best Local Similarity 85.8%; Pred. No. 1.3e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGTCCCTCTCTGGGAGTTGAT 60
DB 257 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGTCCCTCTCTGGGAGTTGAT 316
QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
DB 317 GCCGTGGCAAACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACA 376
QY 121 GTCACCTCTACAGAGCTTCTCTCTGGGCTTTCATCAGCTGGACGGGTACCTGCC 180
DB 377 GTCACCTCTACAGAGCTTCTCTCTGGGCTTTCATCAGCTGGACGGGTACCTGCC 436
QY 181 ATTGTCCAGCCCAACCAAGTCAGTCAAGGCCCAAGGAACTCTGGCTGAAAAGGCTCTAT 240
DB 437 ATCGTCCAGCCCAACCAAGTCAGTCAAGGCCCAAGGAACTCTGGCTGAAAAGGCTCTAT 496
QY 241 GTGGGGTCTGATCCAGCCCTCTCTCTGATATACCTGATCTCTTTGGGAGCTC 300
DB 497 GTGGGGTCTGATCCAGCCCTCTCTCTGATATACCTGATCTCTTTGGGAGCTC 554
QY 301 AGCCAGGGGAGATCAGTCAGGGGAGTACAGTATCTGTGACGGCTTTTACCCCGAT 360
DB 555 -----TCAGTGAGGAGATGACAGATATCTGTGACGGCTTTTACCCCGAT 601
QY 361 AGCCTGTGAGTGGTGTTCATTTCCAGCATATAATGGTGGGTCTCATCTGCGCCGCG 420

DB 602 GACTTGTGGTGGTGTGTGTTTCCAGTTTCCAGCATCATGTTGGCTTATCTCGCTGGT 661
QY 421 ATCGTCATCTCTCTGTTTACTGATCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 480
DB 662 ATTGTTCATCTCTCTGTTTACTGATCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 721
QY 481 CAGAAGCGCAAGCCCTCAAGACGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 540
DB 722 CAGAAGCGCAAGCCCTCAAGACGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 781
QY 541 CTGCCATATTATGTGGGATCAAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 600
DB 782 CTGCCATATTATGTGGGATCAAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 841
QY 601 GGATGTGATCTTCAGAGCATTTGGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 660
DB 842 GGATGTGATCTTCAGAGCATTTGGACAGTCATCTCTTAAGCTGTGACATCTCCAAAGGGCCAC 901
QY 661 TTCCACTGTTGCTGAACCCCATCTCTAT 690
DB 902 TTCCACTGTTGCTGAACCCCATCTCTAT 931

RESULT 6
US-10-305-720-1235
; Sequence 1235, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1235
; LENGTH: 1225
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 GI899313
US-10-305-720-1235

Query Match 75.0%; Score 517.2; DB 16; Length 1225;
Best Local Similarity 85.8%; Pred. No. 1.3e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGTCCCTCTCTGGGAGTTGAT 60
DB 257 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGTCCCTCTCTGGGAGTTGAT 316
QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
DB 317 GCCGTGGCAAACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACA 376
QY 121 GTCACCTCTACAGAGCTTCTCTCTGGGCTTTCATCAGCTGGACGGGTACCTGCC 180
DB 377 GTCACCTCTACAGAGCTTCTCTCTGGGCTTTCATCAGCTGGACGGGTACCTGCC 436
QY 181 ATTGTCCAGCCCAACCAAGTCAGTCAAGGCCCAAGGAACTCTGGCTGAAAAGGCTCTAT 240
DB 437 ATCGTCCAGCCCAACCAAGTCAGTCAAGGCCCAAGGAACTCTGGCTGAAAAGGCTCTAT 496
QY 241 GTGGGGTCTGATCCAGCCCTCTCTCTGATATACCTGATCTCTTTGGGAGCTC 300
DB 497 GTGGGGTCTGATCCAGCCCTCTCTCTGATATACCTGATCTCTTTGGGAGCTC 554
QY 301 AGCCAGGGGAGATCAGTCAGGGGAGTACAGTATCTGTGACGGCTTTTACCCCGAT 360
DB 555 -----TCAGTGAGGAGATGACAGATATCTGTGACGGCTTTTACCCCGAT 601

RESULT 7
 US-10-641-643-1213
 ; Sequence 1213, Application US/10641643
 ; Publication NO. US20040077003A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; Susan G. Stuart
 ; Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 ; OF INVENTION: GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/641,643
 ; FILING DATE: 14-Aug-2003
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: <Unknown>
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1213:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1664 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g219868
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1213 :

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2143
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L06797
US-09-880-107-2143

Query Match 75.0%; Score 517.2; DB 9; Length 1670;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

```
QY 1 CTGACCTGTGAGTGGCTGACCTCCCTTTGTGTCATCACACTCCCTCTCTGGGAGTGTAT 60
DB 304 CTGACCTGTGAGTGGCGGACCTCTCTTTGTGTCATCACACTCCCTCTCTGGGAGTGTAT 363
QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
DB 364 GCCGTGGCAAACTGTGTTGGGAACTTCTATGCAAGGAGTCCATGTCTATCACACA 423
QY 121 GTCAACCTCTACAGCAGGTTCTCATCTGGCTTCTATCAGCTGACCGGTACCTGCC 180
DB 424 GTCAACCTCTACAGCAGGTTCTCATCTGGCTTCTATCAGCTGACCGGTACCTGCC 483
QY 181 ATGTCCACGCCCAACACAGTCAAGGCCAAGGAAACTCTGGCTGAAAGGAGTCTAT 240
DB 484 ATCGTCCACGCCCAACACAGTCAAGGCCAAGGAGTCTGTGGTGAAGGAGTCTAT 543
QY 241 GTGGGCGTCTGGATCCAGCCCTCTCTGACTATATACCTGACCTTTTACCCCGAT 300
DB 544 GTTGGGCGTCTGGATCCCTGCTGCTGACTATATCTGACCTTTTACCCCAAG-- 601
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATATCTGACCTTTTACCCCAAT 360
DB 602 -----TCAGTGAGGAGATGACAGTATATCTGACCTTTTACCCCAAT 648
QY 361 AGCCTGTGATGTGGTGTTCATATCCAGCATATATGTTGGGTCTCTATCTGCGCGG 420
DB 649 GACTTGTGGTGTGGTGTTCAGATTTCCAGATTTCCAGCATATGTTGGCTTATCTGCTGGT 708
QY 421 ATCGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
DB 709 ATTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 768
QY 481 CAGAAGCGCAAGCCCTCAAGACGACAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 829 CTGCTTACTACTACTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
QY 601 GGATGTGACTTCGAGAGCATTTGTCACAGTGGATCTCCATCAGAGGCGCTCTGCTTC 660
DB 889 GGGTGTGAGTTTGAACAACATGTCACAGTGGATTTCCATCAGAGGCGCTCTGCTTC 948
QY 861 TTCCACTGTTGCTGAAACCCCTCTCTAT 690
DB 949 TTCCACTGTTGCTGAAACCCCTCTCTAT 978
```

RESULT 9
US-09-960-706-636
; Sequence 636, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Murgar, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960, 706
; CURRENT FILING DATE: 2001-09-24

; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 636
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 L06797
US-09-960-706-636

```
Query Match 75.0%; Score 517.2; DB 10; Length 1670;  
Best Local Similarity 85.8%; Pred. No. 1.5e-156;  
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
```

QY 1 CTGACCTGTGAGTGGCTGACCTCCCTTTGTGTCATCACACTCCCTCTCTGGGAGTGTAT 60
DB 304 CTGACCTGTGAGTGGCGGACCTCTCTTTGTGTCATCACACTCCCTCTCTGGGAGTGTAT 363
QY 61 GCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
DB 364 GCCGTGGCAAACTGTGTTGGGAACTTCTATGCAAGGAGTCCATGTCTATCACACA 423
QY 121 GTCAACCTCTACAGCAGGTTCTCATCTGGCTTCTATCAGCTGACCGGTACCTGCC 180
DB 424 GTCAACCTCTACAGCAGGTTCTCATCTGGCTTCTATCAGCTGACCGGTACCTGCC 483
QY 181 ATGTCCACGCCCAACACAGTCAAGGCCAAGGAAACTCTGGCTGAAAGGAGTCTAT 240
DB 484 ATCGTCCACGCCCAACACAGTCAAGGCCAAGGAGTCTGTGGTGAAGGAGTCTAT 543
QY 241 GTGGGCGTCTGGATCCAGCCCTCTCTGACTATATACCTGACCTTTTACCCCGAT 300
DB 544 GTTGGGCGTCTGGATCCCTGCTGCTGACTATATCTGACCTTTTACCCCAAG-- 601
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATATCTGACCTTTTACCCCAAT 360
DB 602 -----TCAGTGAGGAGATGACAGTATATCTGACCTTTTACCCCAAT 648
QY 361 AGCCTGTGATGTGGTGTTCATATCCAGCATATATGTTGGGTCTCTATCTGCGCGG 420
DB 649 GACTTGTGGTGTGGTGTTCAGATTTCCAGATTTCCAGCATATGTTGGCTTATCTGCTGGT 708
QY 421 ATCGTATCT 480
DB 709 ATTGTATCT 768
QY 481 CAGAAGCGCAAGCCCTCAAGACGACAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB 769 CAGAAGCGCAAGCCCTCAAGACGACAGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 828
QY 541 CTGCCATATTTATGTTGGGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 600
DB 829 CTGCTTACTACTACTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 888
QY 601 GGATGTGACTTCGAGAGCATTTGTCACAGTGGATCTCCATCAGAGGCGCTCTGCTTC 660
DB 889 GGGTGTGAGTTTGAACAACATGTCACAGTGGATTTCCATCAGAGGCGCTCTGCTTC 948
QY 861 TTCCACTGTTGCTGAAACCCCTCTCTAT 690
DB 949 TTCCACTGTTGCTGAAACCCCTCTCTAT 978

RESULT 10
US-10-372-683-1
; Sequence 1, Application US/10372683
; Publication No. US20040009171A1
; GENERAL INFORMATION:
; APPLICANT: GERRITSEN, MARY E.

```
; APPLICANT: PEALE JR., FRANKLIN V.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: METHODS FOR THE TREATMENT OF CARCINOMA
; FILE REFERENCE: P1928R1P1
; CURRENT APPLICATION NUMBER: US/10/372,683
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 10/271,690
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US 60/344,534
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 1
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-372-683-1

Query Match 75.0%; Score 517.2; DB 16; Length 1670;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGACCTGTGAGTGGCTGACCTCTCTTTGTGTAAGGCTGCTCCATATCATCTCACT 60
DB 304 CTGACCTGTGAGTGGCGACCTCTCTTTGTGTAAGGCTGCTCCATATCATCTCACT 363
QY 61 GCCATGGCTGACTGCTACTTTGGGAAATTTTGTGTAAGGCTGCTCCATATCATCTCACT 120
DB 364 GCCGTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGAGTCCATGCTATACACA 423
QY 121 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCTGAGCGGTACCTCGCC 180
DB 424 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCTGAGCGGTACCTCGCC 483
QY 181 ATTGTCCAGCCACCAAGTCAAGGCGGATGACAGTACATCTGGCTGAAAGGCGAGTCTAT 240
DB 484 ATCGTCCAGCCACCAAGTCAAGGCGGATGACAGTACATCTGGCTGAAAGGCGAGTCTAT 543
QY 241 GTGGGCGTCTGATCCAGCGCTCTCTCTGACTATACCTGACTTCTATCTTTGCCGAGTC 300
DB 544 GTTGGCGTCTGATCCCGCTCTCTCTGACTATTTCCGACTTCTATCTTTGCCAAG-- 601
QY 301 AGCAGGGGACATCAGTCAGGGGATGACAGTACATCTGACCGCTTTTACCCCGAT 360
DB 602 -----TCAGTGGGCGAGTACAGATATATCTGACCGCTTCTATCCCAAT 548
QY 361 AGCTGTGATGGTGGTGTTCATTTCCAGCATATAATGGTGGGTCTCATCTGCCCGGC 420
DB 649 GACTTGTGGTGGTGTGTTCAGTTCAGCATATTCAGCATATTCAGTTCAGCGCTTCTACCCCAAT 548
QY 421 ATCGTCACTCTCTCTGTTACTGATCATCTCTAGCTGTCACATCTCCAGGCGCAC 480
DB 709 ATTGTCACTCTGCTGCTATTTGCAATTCATCTCCAGCTGTCACATCTCCAGGCGCAC 768
QY 481 CAGAAGCGCAAGCGCTCCAGAGCGAGTCACTCTCATCTAGCTTTCTTTGCCCTGCTGG 540
DB 769 CAGAAGCGCAAGCGCTCCAGAGCGAGTCACTCTCATCTAGCTTTCTTTGCCCTGCTGG 828
QY 541 CTGCCATATATATGGGGATCAGCATCTGCTCTCTTCTGAGTCTATCAAGCAA 600
DB 829 CTGCCTTACTACATTGGGATCAGCATCTGCTCTCTTCTATCTCTCTGAAATCATCAAGCAA 888

RESULT 11
US-10-440-464-77
; Sequence 77, Application US/1040464
```

```
; Publication No. US20040018528A1
; GENERAL INFORMATION:
; APPLICANT: DEFRIMO, SAMUEL
; APPLICANT: O'FARRELL, ANNE-MARIE
; FILE REFERENCE: MORIMOTO, ALESSA
; APPLICANT: SMOLICH, BEVERLY
; APPLICANT: KANNING, WILLIAM
; APPLICANT: WALTER, SARAH
; APPLICANT: CHERINGTON, JULIE
; APPLICANT: SCHILLING, JIM
; TITLE OF INVENTION: NOVEL BIOMARKERS OF TYROSINE KINASE INHIBITOR EXPOSURE
; TITLE OF INVENTION: AND ACTIVITY IN MAMMALS
; FILE REFERENCE: 038602/1592
; CURRENT APPLICATION NUMBER: US/10/440,464
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/380,872
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/448,922
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: 60/448,874
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 1670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-440-464-77

Query Match 75.0%; Score 517.2; DB 16; Length 1670;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;

QY 1 CTGACCTGTGAGTGGCTGACCTCTCTTTGTGTAAGGCTGCTCCATATCATCTCACT 60
DB 304 CTGACCTGTGAGTGGCGACCTCTCTTTGTGTAAGGCTGCTCCATATCATCTCACT 363
QY 61 GCCATGGCTGACTGCTACTTTGGGAAATTTTGTGTAAGGCTGCTCCATATCATCTCACT 120
DB 364 GCCGTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGAGTCCATGCTATACACA 423
QY 121 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCTGAGCGGTACCTCGCC 180
DB 424 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCTGAGCGGTACCTCGCC 483
QY 181 ATTGTCCAGCCACCAAGTCAAGGCGGATGACAGTACATCTGGCTGAAAGGCGAGTCTAT 240
DB 484 ATCGTCCAGCCACCAAGTCAAGGCGGATGACAGTACATCTGGCTGAAAGGCGAGTCTAT 543
QY 241 GTGGGCGTCTGATCCAGCGCTCTCTCTGACTATACCTGACTTCTATCTTTGCCGAGTC 300
DB 544 GTTGGCGTCTGATCCCGCTCTCTCTGACTATTTCCGACTTCTATCTTTGCCAAG-- 601
QY 301 AGCAGGGGACATCAGTCAGGGGATGACAGTACATCTGACCGCTTTTACCCCGAT 360
DB 602 -----TCAGTGGGCGAGTACAGATATATCTGACCGCTTCTATCCCAAT 548
QY 361 AGCTGTGATGGTGGTGTTCATTTCCAGCATATAATGGTGGGTCTCATCTGCCCGGC 420
DB 649 GACTTGTGGTGGTGTGTTCAGTTCAGCATATTCAGCATATTCAGTTCAGCGCTTCTACCCCAAT 548
QY 421 ATCGTCACTCTCTCTGTTACTGATCATCTCTAGCTGTCACATCTCCAGGCGCAC 480
DB 709 ATTGTCACTCTGCTGCTATTTGCAATTCATCTCCAGCTGTCACATCTCCAGGCGCAC 768
QY 481 CAGAAGCGCAAGCGCTCCAGAGCGAGTCACTCTCATCTAGCTTTCTTTGCCCTGCTGG 540
DB 769 CAGAAGCGCAAGCGCTCCAGAGCGAGTCACTCTCATCTAGCTTTCTTTGCCCTGCTGG 828
QY 541 CTGCCATATATATGGGGATCAGCATCTGCTCTCTTCTGAGTCTATCAAGCAA 600
DB 829 CTGCCTTACTACATTGGGATCAGCATCTGCTCTCTTCTATCTCTCTGAAATCATCAAGCAA 888
```

QY 601 GGATGTGACTTCGAGAGCATTTGCAAGTGGATCTCCATCAGAGAGCCCTCGCCCTTC 660
Db 889 GGGTGTGAGTTGAGAACACATGTGACAAAGTGGATTTCCATCAGGAGCCCTAGCTTTC 948
QY 661 TTCACACTGTGCTGAAACCCCATCTCTAT 690
Db 949 TTCACACTGTGCTGAAACCCCATCTCTAT 978
RESULT 12
US-10-211-462-80
; Sequence 80, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 80
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-462-80

Query Match 75.0%; Score 517.2; DB 13; Length 1679;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGAGTCCCTCTGCGAGTTGAT 60
Db 320 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGAGTCCCTCTGCGAGTTGAT 379
QY 61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
Db 380 GCCGTGGCAAACTGGTACTTTGGGAACTTCTTATGCAAGGAGTCCATGTATCTACACA 439
QY 121 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCGTGCAGCGGTACCTGCC 180
Db 440 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGTGTGAGCGGTACCTGCC 499
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCCAAGAACTGCTGGCTGAAAGGAGTCTAT 240
Db 500 ATCGTCCAGCCACCAACAGTCAAGGCCCAAGAGTCTGCTGCTGAAAGTGTCTAT 559
QY 241 GTGGCGTGTGGATCCAGCGCTCTCTGATATACCTGATCTGATCTTTCCGAGCGTC 300
Db 560 GTTGGCGTGTGGATCCAGCGCTCTCTGATATACCTGATCTGATCTTTCCGAGCGTC 617
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATCTGTGACCGCTTTACCCCGAT 360
Db 618 -----TCAGTGGGAGATGACAGTATCTGTGACCGCTTTACCCCGAT 664
QY 361 AGCTGTGAGTGTGTGTTTCAATTCAGCATATAATGGTGGTGTCTCATCTGCCCGGC 420
Db 665 GACTTGTGGTGTGTGTGTTTCAATTCAGCATATAATGGTGGTGTCTCATCTGCCCGGC 724
QY 421 ATCGTATCTCTCTCTGTTACTGATCATCATCTCTAAGCTGTACACTCCAGGGCCAC 480

Db 725 ATTGTATCTCTGCTGCTATTGCTATCATCTCAAGCTGTACACTCCAGGAGCCAC 784
QY 481 CAGAAGCGCAAGCCCTCAAGAGCAGTCATCTCATCTAGCTTTCTTTCCCTGCTGG 540
Db 785 CAGAAGCGCAAGCCCTCAAGAGCAGTCATCTCATCTAGCTTTCTTTCCCTGCTGG 844
QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCTCTTCTATCTTTGGGAGTCAATCAAGCA 600
Db 845 CTGCCATTACTATCTGGGATCAGCATCGACTCTCTTCTATCTTTGGGAGTCAATCAAGCA 904
QY 601 GGATGTGACTTCGAGAGCATTTGCAAGTGGATCTCCATCAGAGAGCCCTCGCCCTTC 660
Db 905 GGGTGTGAGTTGAGAACACATGTGCAAGTGGATTTCCATCAGGAGCCCTAGCTTTC 964
QY 661 TTCACACTGTGCTGAAACCCCATCTCTAT 690
Db 965 TTCACACTGTGCTGAAACCCCATCTCTAT 994
RESULT 13
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Dudy, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181,906
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match 75.0%; Score 517.2; DB 13; Length 1679;
Best Local Similarity 85.8%; Pred. No. 1.5e-156;
Matches 592; Conservative 0; Mismatches 83; Indels 15; Gaps 1;
QY 1 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGAGTCCCTCTGCGAGTTGAT 60
Db 320 CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCATCAGAGTCCCTCTGCGAGTTGAT 379
QY 61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
Db 380 GCCGTGGCAAACTGGTACTTTGGGAACTTCTTATGCAAGGAGTCCATGTATCTACACA 439
QY 121 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGCGTGCAGCGGTACCTGCC 180
Db 440 GTCAACCTCTACAGAGCGTTCTCATCTGGCTTCATCAGTGTGAGCGGTACCTGCC 499
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCCAAGAACTGCTGGCTGAAAGGAGTCTAT 240
Db 500 ATCGTCCAGCCACCAACAGTCAAGGCCCAAGAGTCTGCTGCTGAAAGTGTCTAT 559
QY 241 GTGGCGTGTGGATCCAGCGCTCTCTGATATACCTGATCTGATCTTTCCGAGCGTC 300
Db 560 GTTGGCGTGTGGATCCAGCGCTCTCTGATATACCTGATCTGATCTTTCCGAGCGTC 617
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATCTGTGACCGCTTTACCCCGAT 360
Db 618 -----TCAGTGGGAGATGACAGTATCTGTGACCGCTTTACCCCGAT 664
QY 361 AGCTGTGAGTGTGTGTTTCAATTCAGCATATAATGGTGGTGTCTCATCTGCCCGGC 420


```
QY 181 ATTCTCCACGCCACCAACAGTCAAGGCCAAGGAAACTGTGGCTGAAAGGCGAGTCTAT 240
  |||
Db 500 ATCGTCCACGCCACCAACAGTCAAGGCCAAGGAAAGCTGTGGCTGAAAGGCTGCTAT 559
  |||
QY 241 GTGGCGCTGTGATCCACAGCCCTCTCTGACTATACCTGACTTCATCTTTGGCGAGTC 300
  |||
Db 560 GTTGGCGCTGTGATCCCTGCGCCCTCTCTGACTATCCCGACTTCATCTTTGCCAAG-- 617
  |||
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT 360
  |||
Db 618 -----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 664
  |||
QY 361 AGCCTGTGGATGGTGGTGTTCATATCCAGCATATATATGATGGGTCTCATCCTGCGCGGC 420
  |||
Db 665 GACTTGTGGGTGGTGTGTTCAGTTTCAGTTTCAGCACATCATGGTTGGCCTTATCCTGCTGT 724
  |||
QY 421 ATCGTCATCCTCTCTCTGTACTGTGATCATCATCTCTTAAGCTGTCACTCCAAGGGCCAC 480
  |||
Db 725 ATTGTATCCTGTCCTCTCTATTTGATTTATCATCTCAAGCTGTCACTCCAAGGGCCAC 784
  |||
QY 481 CAGAAGCGCAAGGCCCTCAAGACACAGTCACTCTCATCTCTAGCTTCTTTGGCTGCTGG 540
  |||
Db 785 CAGAAGCGCAAGGCCCTCAAGACACAGTCACTCTCATCTCTGCTTCTTTGGCTGCTGG 844
  |||
QY 541 CTGCCATATTATGTGGGATCAGCATCGACTCCTTCATCCTTTTGGAGTCATCAAGCAA 600
  |||
Db 845 CTGCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCTCTGGAATCATCAAGCAA 904
  |||
QY 601 GGATGTGACTTCGAGACATTGTGACAAAGTGGATCTCCATCACAGAGGCCCTCGCCTTC 660
  |||
Db 905 GGGTGTGAGTTTGAGAACACTGTGACACAGTGGATTTCCATCACCGAGGCCCTAGCTTTC 964
  |||
QY 661 TTCCACTGTTGCTGAAACCCCTCTCTAT 690
  |||
Db 965 TTCCACTGTTGCTGAAACCCCTCTCTAT 994
  |||
```

Search completed: May 17, 2004, 22:14:04
Job time : 341.746 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:50:31 ; Search time 1991.6 Seconds
(without alignments)
10345.919 Million cell updates/sec

Title: US-09-367-052-3

Perfect score: 690

Sequence: 1 ctgcacctgtcagtgctga.....gcttgacccatctcttat 690

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hcc.*

9: gb_estl.*

10: gb_est2.*

11: gb_hcc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_eston.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517.2	75.0	872	14	CD251331
2	511.6	74.1	1201	13	CD251287
3	506.2	73.4	1201	13	CD251287
4	505.4	73.2	1201	13	CD251287

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	494.8	71.7	1201	13	CD251331
6	492.2	71.3	1091	13	CD251331
7	491.8	71.3	1201	13	CD251331
8	484	70.1	1051	12	CD251331
9	476.4	69.0	1201	13	CD251331
10	468	67.8	906	13	CD251331
11	453.8	65.8	1201	13	CD251331
12	436.8	63.3	1201	13	CD251331
13	435.6	63.1	749	14	CD251331
14	435	63.0	582	12	CD251331
15	431.8	62.6	952	13	CD251331
16	430.6	62.4	817	14	CD251331
17	419.6	60.8	782	14	CD251331
18	419	60.7	846	10	CD251331
19	418	60.6	446	9	CD251331
20	411.2	59.6	945	13	CD251331
21	409.4	59.3	750	12	CD251331
22	409	58.3	911	13	CD251331
23	404.6	58.6	910	12	CD251331
24	397.8	57.7	891	12	CD251331
25	397.2	57.6	793	9	CD251331
26	387.4	56.1	886	13	CD251331
27	386	55.9	950	13	CD251331
28	386	55.9	612	12	CD251331
29	376.4	54.6	528	12	CD251331
30	376.2	54.5	553	12	CD251331
31	375.6	54.4	928	14	CD251331
32	370	53.6	827	12	CD251331
33	368.4	53.4	1034	12	CD251331
34	366	53.0	600	13	CD251331
35	362.2	52.5	765	12	CD251331
36	361.8	52.4	876	14	CD251331
37	359	52.0	639	9	CD251331
38	354.8	51.4	532	10	CD251331
39	339	49.1	875	12	CD251331
40	334.2	48.4	499	10	CD251331
41	333.2	48.3	581	14	CD251331
42	331.2	48.0	976	13	CD251331
43	328.4	47.6	958	12	CD251331
44	324.8	47.1	444	12	CD251331
45	323.8	46.9	657	13	CD251331

ALIGNMENTS

RESULT 1
CD251331
LOCUS
DEFINITION AGENCOURT 14211919 NIH MGC 179 Homo sapiens cDNA clone
IMAGE:303E4274 5', mRNA sequence.
ACCESSION CD251331
VERSION CD251331.1 GI:31011797
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 872)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaapbs-r@mail.nih.gov
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (ILLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at:

CD251331 872 bp mRNA linear EST 22-MAY-2003

IMAGE:303E4274 5', mRNA sequence.

ACCESSION CD251331

VERSION CD251331.1 GI:31011797

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 872)

NIH-MGC http://mgc.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Daniela S. Gerhard, Ph.D.

COMMENT Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgaapbs-r@mail.nih.gov

```

http://image.llnl.gov
Plate: ND4M449 row: h column: 11
High quality sequence stop: 654.
Location/Qualifiers
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30384274"
/tissue_type="pituitary"
/lab_host="DH10B-iron A ( T1 and T5 phage resistances )"
/clone_lib="NIH_MGC_179"
/note="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV
(destroyed); Site 2: NotI; Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.1 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

```

ORIGIN

Query Match	75.0%	Score 517.2;	DB 14;	Length 872;
Best Local Similarity	85.8%	Pred. No. 2.7e-11;		
Matches 592;	Conservative	0;	Mismatches 83;	Indels 15;
				Gaps 1;

1	QY	CTGCACTGTGCAGTGGCTGACCTCCTCTTTGTGTATACACTCCCTTCTGGGAGTGTGAT	60
59	Db	CTGCACTGTGCAGTGGCGGACCTCCTCTTTGTATCAGCGTTCCTTCTGGGACAGTGTGAT	118
61	QY	GCATGGCTGTACTGTACTTTTGGAAATTTTGTGTAAAGCTGTCCATATCATCTACACT	120
119	Db	GCGTGGCAACTGTGTACTTTTGGAACTTCTATAGCAGCAGTCAATGTATCTACACA	178
121	QY	GTCAACCTCTTACAGCAGCGTTCTCATCCTGGCTTTCATCAGCCTGACCGGTACTCTGCC	180
179	Db	GTCAACCTCTTACAGCAGTGTCTCATCCTGGCTTTCATCAGTCTGACCGCTACCTGGCC	238
181	QY	ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAATGCTGTGGCTGAAAGGACCTCAT	240
239	Db	ATCGTCCAGCCACCAACAGTCAAGGCCAAGGAGCTGTTGGCTGAAAGGTGTCTAT	298
241	QY	GTGGGCGTCTGGATCCCAAGCCCTCCTCTCGTCACTATACCTGACTTCATCTTTGCCGACGTC	300
299	Db	GTTGGCGTCTGGATCCCTGGCCCTCCTGCTGACTATTCCGACTTCATCTTTGCCAAGC	356
301	QY	AGCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGAT	360
357	Db	-----TCAGTGAGGCGAGATGCAGATATATCTGTGACCGCTTCTACCCCAAT	403
361	QY	AGCCTGTGGATGGTGGTGTTCATATCCAGCATATAATGGTGGGTCTCATCTGCGCCGCG	420
404	Db	GACTTGTGGTGGTGTGTTCAGGTTTCAGCATCATGTTGGCTTATCTGCTGGT	463
421	QY	ATCGTCACTCTCTCTGTATTGTCATCATCTTAAGCTGTCACTCCCAAGGGCCAC	480
464	Db	ATTGTCACTCTCTCTGTATTGTGATTTATCATCTCCAAGTGTGCATCTCAAGGGCCAC	523
481	QY	CAGAAGCGCAAGGCCCTCAAGACACAGTTCATCTCATCTAGCTTCTTTGCCCTGTGG	540
524	Db	CAGAAGCGCAAGGCCCTCAAGACACAGTTCATCTCATCTGGCTTCTTCGCTGTGG	583
541	QY	CTGCCATATTATGTGGGATCAGCATCGACTCCTTTCACTCTTTGGAGTTCATCAAGCAA	600
584	Db	CTGCCCTTACTACATTGGGATCAGGATCGACTCCTTTCATCTCTCGGAAATCATCAAGCAA	643
601	QY	GGATGTGACTTCAGAGGATTTGCACAAGTGGATCTCCATCAGAGGCCCTCGCCCTTC	660
644	Db	GGGTGTGAGTTTGAGAACACTGTGCAACAGTGGATTCCATCACCAGGCCCTTACTGTTTC	703
661	QY	TTCCACTGTGCTGAACCCCATCCTCAT	690
704	Db	TTTCCACTGTGTCTGAACCCCATCCTCAT	733

RESULT 2	ACCESSION	REFERENCE
EX421287	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

FEATURES

```

1. 12901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG004YG09"
/tissue_type="B CELLS (R
/cell_line="RAMOS CELL L
/clone_lib="Homo sapiens
/notes="Vector: pcwv5pork
with a NotI-oligo(dT) pri
double-strand cDNA was d
the Not I and EcoRV site
Library was not normaliz

```

ORIGIN .

Query Match	74.1%;	Score 511.6;	DB 13;	Length 1201;
Best Local Similarity	84.6%;	Pred. No. 6.6e-110;		
Matches 584;	Conservative 6;	Mismatches 85;	Indels 15;	Gaps 1;
QY	1	CTGCACCTGTCAAGTGACCTCCCTTTGTTCATCACACTCCCTCTGGGCAGTTGAT	60	
Db	334	CTGCACCTGTCAAGTGCCACCTCCCTTTGTTCATCACGCTTCCCTCTGGGCAGTTGAT	393	
QY	61	GCCATGGCTGACGTGGTACTTTTGGGAAATTTTGTGTAAGCGCTGTCCATATCATCTACACT	120	
Db	394	GCCGTGGCAAACTGGTACTTTTGGGAATTCCTATGTCGAAGCAGTCCATGTCTATCTACACA	453	
QY	121	GTCAACCTCTACAGCAGCGTTTCTCATCTCGGCGCTTCATCAGCCTGCAGCCGTAACCTCGCC	180	
Db	454	GTCAACCTCTACAGCAGTGTCCCTCATCTCGGCGCTTCATCAGTCTGCAGCCGTAACCTCGCC	513	
QY	181	ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTCAAAAGGCAGTCTAT	240	
Db	514	ATCGTCCAAGCCACCAACAGTCAAGGCCAAGGAACTGTTGGCTGAAAAGGTGGTCTAT	573	
QY	241	GTGGCGGTCTGGATCCCAAGCCCTTCCTGCACTATACCTGACTTTCATCTTTGCCGACGTC	300	
Db	574	GTTGGCGTCTGGATCCCTGCGCTTCTGTGCTACTATTCCCGACTTTCATCTTTGCCAAG--	631	
QY	301	AGCCAGGGGACATTCAGTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGAT	360	
Db	632	-----TCAGTAAGCCAGATGACAGATATATCTGTGACCGCTTCTTACCCCAAT	678	
QY	361	AGCCTCTGGATGGTGTGTTTCAAATTCACGATATATATGGTGGGTCTCATCTCTGCCCGC	420	

679	DB	GACTTGTGGGTGGTGTGTTCTCCAGTTTCAGACATCATGGTTGGCCTTATCTCTGCCTGGT	738
421	QY	ATGCTCATCTCTCTCTCTGTTACTGCATCATCATCTCTTAAGCTGTCTACATCCAAAGGGCCAC	480
739	DB	ATTGTGATCTCTGTCTCTGCTATTTCATTATCATCTCCAGCTGTCTACATCCAAAGGGCCAC	798
481	QY	CAGAAGCGCAGAGGCCCTCAAGACGACAGTCATCCTCATCTAGCTTTCTTTGGCTGTGCTGG	540
739	DB	CAGAAGCGCAGAGGCCCTCAAGACCAAGTCATCTCTATCTGTGGCTTTCTTGCCTGTGTGG	858
541	QY	CTGCCATATTATGTGGGGATCAGCATCGACTCTCTTCATCTTTTGGGAGTCATCAAGCAA	600
859	DB	CTGCTTACTACATTGGGGATCAGCATCGACTCTCTTCATCTCTCGGAATCATCAAGCAA	918
601	QY	GGATGTGACTTTCGAGAGCATTTGTGCACAAGTGGATCTCCATCACAAGAGCCCTCGCCTTC	660
919	DB	GGGTGTGATTTTGAAACAATGTGACACAGTGGGATTTCATCCCGAGGCCCTAGCTTTT	978
661	QY	TTCCCACTGTTCCTGGAACCCCACTCTCTAT	690
979	DB	TTTCACCTGTGTGTGAACCCCTCTCTAT	1008

RESULT 3	1201 bp	mrna	linear	EST 05-MAY-2003
EX358346	EX358346	Homo sapiens	PLACENTA	COT 25-NORMALIZED
LOCUS	EX358346	Homo sapiens	PLACENTA	COT 25-NORMALIZED
DEFINITION	clone CS01036YJ17	5-PRIME, mrna sequence.		

EX358346.1	GI:30370215	
EST.		
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1201)	
AUTHORS	Li, W.B., Gruber, C., Jesses, J. and Polayes, D.	
TITLE	Full-length cDNA libraries and normalization	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Genoscope Genoscope, Centre National de Sequencage BP 191 91006 Evry cedex - France Email: segre@genoscope.cns.fr , Web : www.genoscope.cns.fr Library was constructed by life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0D1036CE090P1.	

```
FEATURES
source
1..1201
location/Qualifiers
tataaa_yrrenda_zenobioscye_sequence_id : C6D0A030C6D0F9F.r.
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="CSQDI036XJ17"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 73.4%; Score 506.2; DB 13; Length 1201;
Best Local Similarity 85.3%; Pred. No. 1.2e-108;
Matches 580; Conservative 1; Mismatches 85; Indels 14; Gaps 1;

1	CTGCACTGTGAGTGGCTGACCTCTCTTTGTGATCACA	CTCCCTTCTGGGCACTGAT	60
364	CTGCACTGTGAGTGGCGACCTCTCTTTGTGATCACA	CTCCCTTCTGGGCACTGAT	423
61	GCATGGCTGACTGGTACTTTTGGGAAATTTTGTGAAGGCTGT	CCATATCATCTACACT	120
424	GCCTGGGAAACTGGTACTTTTGGAACTTCCTATGCAAGGCACT	CATGTCACTACACA	483

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D007YML2"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 73.2%; Score 505.4; DB 13; Length 1201;
Best Local Similarity 85.0%; Pred. No. 1.9e-108;
Matches 578; Conservative 3; Mismatches 85; Indels 14; Gaps 1;

QY 1 CTGCACCTCTCAGTGGCTGACCTCTCTTTGTCATCACACATCCCTCTTGGGCGATTGAT 60
DB |||||
QY 358 CTGCACCTCTCAGTGGCGACCTCTCTTTGTCATCACAGCTTCCCTTCTGGGCGATTGAT 417
DB |||||
QY 61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
DB |||||
QY 418 GCCGTGGCAAACTGGTACTTTGGGAATCTCTATGCMAGGAGTCCATGTCTATCACACA 477
DB |||||
QY 121 GTCAACCTCTCAGCAGCGCTTCTATCTCTGGCTTCATCAGCTGAGCGGTACTCTGCC 180
DB |||||
QY 478 GTCAACCTCTCAGCGGTCTCTCTCTCTGGCTTCATCAGTCTGAGCGGTACTCTGCC 537
DB |||||
QY 181 ATTGTCCAGCCACCAACAGTCAAGAGGCAAGAAAGTCTGGCTGAAAGGCGATCTAT 240
DB |||||
QY 538 ATCGTCCAGCCACCAACAGTCAAGAGGCAAGAGTCTGGCTGAAAGGCGATCTAT 597
DB |||||
QY 241 GTGGGCGTCTGATGCCAGCCCTCTCTGATATACCTGATTCATCTTTGCCAGCTC 300
DB |||||
QY 598 GTTGGCGTCTGATGCCAGCCCTCTCTGATATACCTGATTCATCTTTGCCAGCTC 657
DB |||||
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 360
DB |||||
QY 658 AGTGA-----NGGAGATGACAGATATCTGTGACCGCTTTACCCCGAT 703
DB |||||
QY 361 AGCCTGTGATGGTGGTGTTCAAATCCAGCATATATGGTGGGCTCATCTGCCCGC 420
DB |||||
QY 704 GACTTGTGGGTGGTGTGTTCCAGTTTCAGCACATCATGTTGGCTTATCTGCTGCT 763
DB |||||
QY 421 ATCGTCATCTCTCTGTTACTGTCATCATCTCTAGCTGTCACTCCAGGGCCAC 480
DB |||||
QY 764 ATTGTATCTCTCTCTGTTATGCAATATCATCTCAAGCTGTCACTCCAGGGCCAC 823
DB |||||
QY 481 CAGAAGCGCAAGCCCTCAAGACGACAGTCACTCTCATCTAGCTTTCTTTGCCCTGCTGG 540
DB |||||
QY 824 CAGAAGCGCAAGCCCTCAAGACGACAGTCACTCTCATCTAGCTTTCTTTGCCCTGTTGG 883
DB |||||
QY 541 CTGCCATATATGTGGGGATCAGCATCGACTCTCTCTCATCTCTTTTGGGAGTCAATCAAGCAA 600
DB |||||
QY 884 CTGCCCTTACTACATTGGGATCAGCATCGACTCTCTCTCATCTCTCTGGAATCATCAAGCAA 943
DB |||||
QY 601 GGATGTGACTTCGAGAGCATTTGTGACAGTGGATCTCCATCAGAGGCGCTTGCCTTC 660
DB |||||
QY 944 GGGTGTGAGTTGAGAACACTGTGACAAAGTGAGTTTCCATCAGGCGCTTAGCTTTC 1003
DB |||||
QY 661 TTCCACTGTTGCTGAACCC 680
DB |||||
QY 1004 TTCCACTGTTGCTGAACCC 1023
DB |||||

RESULT 5

BX396365 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX396365 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D015YL20 5-PRIME, mRNA sequence.
ACCESSION BX396365
VERSION BX396365.1 GI:30609085

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1272.r For more information about this cluster, see
http://www.genoscope.cns.fr/
csi-bin/cluster.cgi?seq=CS0D015YDF100P1&cluster=1272.r. Contact :
Peng Liang Email : filiang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D015YDF100P1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D015YL20"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 71.7%; Score 494.8; DB 13; Length 1201;
Best Local Similarity 85.3%; Pred. No. 6.1e-106;
Matches 580; Conservative 1; Mismatches 83; Indels 16; Gaps 2;

QY 1 CTGCACCTCTCAGTGGCTGACCTCTCTTTGTCATCACACATCCCTCTTGGGCGATTGAT 60
DB |||||
QY 350 CTGCACCTCTCAGTGGCGACCTCTCTTTGTCATCACAGCTTCCCTTCTGGGCGATTGAT 409
DB |||||
QY 61 GCCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGGCTGTCCATATCATCTACACT 120
DB |||||
QY 410 GCCGTGGCAAACTGGTACTTTGGGAATCTCTATGCAAGGAGTCCATGTCTATCACACA 469
DB |||||
QY 121 GTCAACCTCTCAGCAGCGTTCATCTCTGGCTTCATCAGCTGAGCGGTACTCTGCC 180
DB |||||
QY 470 GTCAACCTCTCAGCAGCTGTCTCTCTCTGGCTTCATCAGTCTGAGCGGTACTCTGCC 529
DB |||||
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCAAGAAAGTCTGGCTGAAAGGCGAGTCTAT 240
DB |||||
QY 530 ATCGTCCAGCCACCAACAGTCAAGGCAAGAGCTGTGGCTGAAAGGCTGTCTAT 589
DB |||||
QY 241 GTGGGCGTCTGGATCCAGCGCTCTCTCTGATATACCTGATTCATCTTTGCCAGCTC 300
DB |||||
QY 590 GTTGGGCTCTGGATCCCTCTCTCTGCTGACTATTCCGAGTTCATCTTTGCCAAG-- 647
DB |||||
QY 301 AGCCAGGGGACATCAGTCAAGGGGATGACAGTATCTGTGACCGCTTTACCCCGAT 360
DB |||||
QY 648 -----TCAGTAGGGCAGATGACAGATATCTGTGACCGCTTTACCCCAAT 694
DB |||||
QY 361 AGCCTGTGATGGTGGTGTTCAAATCCAGCATATATGGTGGGCTCTCATCTCTGCCCGC 420
DB |||||
QY 695 GACTTGTGGGTGGTGTGTTCCAGTTCAGCACATCATGTTGGCTTATCTCTGCTGCT 754
DB |||||
QY 421 ATCGTATCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTCACTCCAGGGCCAC 480
DB |||||
QY 755 ATTGTCTCTCTCTCTGCTTATGTCATTTATCTCTCAAGCTGTCACTCCAGGGCCAC 814
DB |||||
QY 481 CAGAAGCGCAAGCCCTCAAGAGCAGCATCTCTCATCTCTAGCTTTCTTTGCTCTCTGG 540
DB |||||

```
Db      815  CAGAAGCGCAAGCCCTCAAGACCACAGTCATCCTCCTGCTTTCTTGGCTGTTGG 874
Qy      541  CTGCGATATATATGTTGGGATCAGATCGACTCTTCTCATCTCTTTTGGAGTATCAAGCAA 600
Db      875  CTGCGTACTACATTTGGGATCAGATCGACTCTTCTCATCTCTTGGAAATCATCAAGCAA 934
Qy      601  GGATGTGACTTTCAGAGCACTTTGGCAAGAGTGATCTCCATCAGAGGCGCTCGGCTTC 660
Db      935  GGGTGTGAGTTTGAGAACACTGTGCACAGTGGAT-TYCATCACCGAGGCGCTAGCTTTC 993
Qy      661  TTCCACTGTGCTGCTGAACCC 680
Db      994  TTCCTACTGTGTCTGAACCC 1013

RESULT 6
BX397685
LOCUS
DEFINITION
BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1041Y013 5-PRIME, mRNA sequence.
ACCESSION
BX397685
VERSION
BX397685.1 GI:30621359
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1091)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1041AH07QP1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1041AH07QP1.
FEATURES
source
1..1091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1041Y013"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 71.3%; Score 492.2; DB 13; Length 1091;
Best Local Similarity 85.0%; Pred. No. 2.4e-105;
Matches 588; Conservative 2; Mismatches 86; Indels 16; Gaps 3;
Qy      1  CTGCACTGTGAGTGTGACCTCTCTTGTGTCATCAGACTCCCTCTTGGGAGTGTGAT 60
Db      352  CTGCACTGTGAGTGTGACCTCTCTTGTGTCATCAGACTCCCTCTTGGGAGTGTGAT 411
Qy      61  GCATGGCTGACGTGGTACTTTGGGAAATTTTGTGTAGAGCTGTGTCATCATCTACACT 120
Db      412  GCGTGGCAAACTGGTACTTTGGGAAATTTTGTGTAGAGCTGTGTCATCATCTACACA 471
Qy      121  GTCAACTCTTACAGCAGCGTGTCTCATCTGGGCTTTCATCAGCTGGACGGTACCTCGCC 180
Db      472  GTCAACTCTTACAGCAGGTGTCTCATCTGGGCTTTCATCAGCTGGACGGTACCTCGCC 531
```

```
Qy      181  ATTGTCCAGCCCAACACAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTCTAT 240
Db      532  ATCTGCCAGCGCCACCAACAGTCAAGGCCAAGAACTGCTGGCTGAAAGGAGTCTAT 591
Qy      241  GTGGGGCTGTGGATCCAGAGCCCTCTCTCTGACTATATACCTGACTTTCATCTTTGCGAGGTC 300
Db      592  GTTGGGGCTGTGGATCCCTGCTCTCTGACTATTTCCCGACTTCACTTTGCGCAACGTC 651
Qy      301  AGCCAGGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
Db      652  AGTGA-----NCGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 697
Qy      361  AGCCTGTGGATGGTGGTGTTCATTTCCAGCATATATGTTGGGTCTCATCTCTCCCGGCG 420
Db      698  GACTTGTGGGTGGTGGTGTTCAGAGTTCAGACATCATGTTGGCTTATCTTGCCTGGT 757
Qy      421  ATCTCATCTCTCTCTCTTACTGTCATCATCTTCTTAAGCTGTGTCACATCAAGGGCAC 480
Db      758  ATTCTCATCTCTCTCTCTTATGCAATATATCTCTCAAGCTGTGTCACATCAAGGGCAC 817
Qy      481  CAGAAGCGCAAGGCCCTCAAGACGACAGTCTATCTCATCT-AGCTTTCTTTGGCTGCT- 538
Db      818  CAGAAGCGCAAGGCCCTCAAGACGACAGTCTATCTCATCTCTGCTTTCTTGGCTGTTK 877
Qy      539  GGCTGCCATATTATGTGGGATCAGCATCGACTCTCTTCTCATCTCTTTGGGAGTCAACAAGC 598
Db      878  GGCTGCCATATTATGTGGGATCAGCATCGACTCTCTTCTCATCTCTTGGGAGTCAACAAGC 937
Qy      599  AAGCATGTGATTCGACAGCATTTGTGCAAGTGGATCTCCATCAGAGGCGCTCGCCT 658
Db      938  AAGGTGTGAGTTTGAGAACACAGTGTGCAAGTGGATTTCCATCAGGAGCGCTAGCTT 997
Qy      659  TCTTCCACTGTGCTGAACCCCATCTCTAT 690
Db      998  TTTTCCACTGTGTTTGAACCCCATCTCTAT 1029

RESULT 7
BX353253
LOCUS
DEFINITION
BX353253 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC003YG03 5-PRIME, mRNA sequence.
ACCESSION
BX353253
VERSION
BX353253.1 GI:30347726
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC003AD02QP1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DC003AD02QP1.
FEATURES
source
1..1201
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC003YG03"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
```

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN	
Query Match	71.3%; Score 491.8; DB 13; Length 1201;
Best Local Similarity	84.9%; Pred. No. 3.1e-105;
Matches 577; Conservative	2; Mismatches 85; Indels 16; Gaps 2;
QY	1 CTGCACCTGTGAGTGGTGGTACCTCTCTTTGTCATCACATCCCTCTCTGGGAGTTGAT 60
DB	367 CTGCACCTGTGAGTGGGCGACCTCTCTTTGTCATCACGCTTCTCTGGGAGTTGAT 426
QY	61 GCCATGGCTGACTGTAATTTGGGAAATTTTGTAAAGGTGTCATATCATCTACACT 120
DB	427 GCCGTGGCAAACTGTAATTTGGGAACTTCTATGCAAGGCGAGTCCATGTACACACA 486
QY	121 GTCACCTCTACAGCAGGTTCTCATCTGCGCTTCATCGCCGTCGACCGTACCTGCC 180
DB	487 GTCACCTCTACAGCAGGTTCTCATCTGCGCTTCATCGCTGCGACCGTACCTGCC 546
QY	181 ATTGTCCACGCGCACCAACAGTCAGGCGCAAGGAACTGCTGGCTGAAAGGCGAGTCTAT 240
DB	547 ATCGTCCAGCGCACCAACAGTYAGAGGCGCAAGGAGCTGTGGCTGAAAGGTGCTAT 606
QY	241 GTGGCGTCTGATCCAGCCCTCTCTGCTGACTATACCTGACTGACTCTTGGCGAGTC 300
DB	607 GTTGGCGTCTGATCCAGCCCTCTCTGCTGACTATCCCGACTTCATCTTTGCCAAGC-- 664
QY	301 AGCCAGGGGACATCAGTCAGGCGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
DB	665 -----TCAGTANGCGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 711
QY	361 AGCCTGTGGATGGTGTGTTCAATTCAGCATATATATGTTGGTGTCTATCTCTGCCGCG 420
DB	712 GACTTGTGGTGGTGTGTTCCAGTTTCAGACATCATGTTGGTGGCTTATCTCTGCTGGT 771
QY	421 ATCGTCATCTCTCTCTGTTACTGATCATCTCTAAGCTGTGACACTCCCAAGGGCCAC 480
DB	772 ATGTGATCTCTCTCTGTTGATATATCTCTAAGCTGTGACACTCCCAAGGGCCAC 831
QY	481 CAGAAGCGCAAGGCGCTCAAGACAGCATCTCTAAGCTGTGACACTCCCAAGGGCCAC 540
DB	832 CAGAAGCGCAAGGCGCTCAAGACACAGTCATCTCTCTCTGCTGCTTCTCTCGCTGKGG 891
QY	541 CTGCCATATTATGTTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB	892 CTGCCCTTACTACTTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 951
QY	601 GGATGTGACTTCGAGAGCATTTGTGACAAAGTGGATCTCCATCAGAGGCGCTCTGCTTC 660
DB	952 GGGTGTGAGTTTGAGAACACTGTGACAAAGTGGATTTCCATCAGGAGGCGCTCTCTCT 1011
QY	661 TTCACCTGTGCTGAAACC 680
DB	1012 TT-CACCTGTGCTGAAACC 1030

RESULT 8
BM545259
LOCUS
DEFINITION
AGENCY 6497171 NIH_MGC_124 Homo sapiens cdna clone IMAGE:5726963
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1051)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12720 row: e column: 12
High quality sequence start: 20
High quality sequence stop: 753.
Location/Qualifiers
1. .1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5726963"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV
(destroyed); Site 2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

FEATURES

Source

ORIGIN

Query Match	70.1%; Score 484; DB 12; Length 1051;
Best Local Similarity	84.5%; Pred. No. 2e-103;
Matches 585; Conservative	0; Mismatches 90; Indels 17; Gaps 3;
QY	1 CTGCACCTGTGAGTGGTGGTACCTCTCTTTGTCATCACATCCCTCTCTGGGAGTTGAT 60
DB	299 CTGCACCTGTGAGTGGGCGACCTCTCTTTGTCATCACGCTCTCTCTGGGAGTTGAT 358
QY	61 GCATGGCTGACTGTGTAATTTGGGAAAATTTTGTGTAAAGCTGTCCATATCATCTACACT 120
DB	359 GCGGTGGCAAACTGTTGTTGGGAACTTCTATGCAAGGCGAGTCCATGTCTATACACA 418
QY	121 GTCAACCTCTACAGCAGGCTTCTCATCTGGCTTTCATCAGCTCGACCGGTACCTCGCC 180
DB	419 GTCAACCTCTACAGCAGGCTTCTCATCTGGCTTTCATCAGCTCGACCGGTACCTCGCC 478
QY	181 ATTGTCCAGCCCAACAGTCATCAAGGCGCAAGGAACTGCTGGCTGAAAAGGAGTCTAT 240
DB	479 ATCGTCCAGCCCAACAGTCATCAAGGCGCAAGGAACTGCTGGCTGAAAAGGAGTCTAT 538
QY	241 GTGGGCTGTGGATCCAGCCCTCTCTGCTGACTATACCTGACTTCTCTTTGGCGAGCTC 300
DB	539 GTTGGGCTGTGGATCCAGCCCTCTCTGCTGACTATCCCGACTTCTCTTTGGCGAGCTC 596
QY	301 AGCCAGGGGACATCAGTCAGGCGGATGACAGGTACATCTGTGACCGCTTTACCCCGAT 360
DB	597 -----TCAGTGAGGCGAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 643
QY	361 AGCCTGTGGATGGTGTGTTTCAATTCAGCATATATATGTTGGTGTCTATCTCTGCCGCG 420
DB	644 GACTTGTGGTGGTGTGTTTCCAGTTTCAGGACATCATGTTGGCTTATCTCTGCTGCTGT 703
QY	421 ATCGTCATCTCTCTCTGTTTACTGTCATCATCTCTTAAGCTGTCTACACTTCAAGGGCCAC 480
DB	704 ATGTGATCTCTCTCTGCTTATGCTATTCATCTTCAAGCTGTCTACATCCAGGGCCAC 763
QY	481 CAGAAGCGCAAGGCGCTCAAGACGACAGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB	764 CAGAAGCGCAAGGCGCTCAAGACCAAGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823

QY	541	CTGCGATATTATGTGGGGATCAGCATCGACTCTTTCATCTCTTTTGGGAGTCATCAAGCAA	600
Db	824	CTGCGCTTACTACATTTGGGATCAGCATCGACTCTTTCATCTCTCTGGAATCATCAAGCAA	883
QY	601	-GGATGTGACTTTCAGAGCATTGTGCACAGTGGATCTCCATCACAGA-GGCCCTCCCT	658
Db	884	GGGTTGTAGTTTGAGAACACGGTGACAGTGGATTTTCATCACCAGGGCCCTAGCTT	943
QY	659	TCTTCCACTGTTGCCTGAACCCCATCTCTAT	690
Db	944	TCTTCACTGTTGGCTGAACCCCATCTCTAT	975
RESULT 9			
LOCUS	BX462627	1201 bp	linear EST 22-MAY-2003
DEFINITION	BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA clone CS0DH005YK10 5-PRIME, mRNA sequence.		
ACCESSION	BX462627		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1272.r For more information about this cluster, see http://www.genoscope.cns.fr/ cqi-bin/cluster.cgi?seq=CS0DH005BF05QF1&cluster=1272.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DH005BF05QF1.		
FEATURES			
source	1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DH005YK10" /tissue_type="T CELLS (JURKAT CELL LINE)" /cell_line="JURKAT CELL LINE" /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)" /note="vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned, into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN			
Query Match	69.0%	Score 476.4;	DB 13; Length 1201;
Best Local Similarity	84.7%	Pred. No. 1.3e-101;	
Matches	576;	Conservative 0; Mismatches 86; Indels 18; Gaps 3;	
QY	1	CTGCACCTGTGAGTGGCTGACCTCTCTTTGTCTATCATCATCCCTTCTGGGAGTTGAT	60
Db	364	CTGCACCTGTGAGTGGCGGACCTCTCTTTGTCTATCATCATCCCTTCTGGGAGTTGAT	423
QY	61	GCCATGGCTGACTGGTACTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACAT	120
Db	424	GCGGTGGCAACTGGTACTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACAT	483
QY	121	GTCAACCTCTACAGCAGCGTTTCTCATCTGGCCCTTCATCGCCTGGACCGTACCTGCC	180
Db	484	GTCAACCTCTACAGCAGTGTCTCTCTGCTTTCATGCTGACCGCTACCTGCTGCTG	543

/note="Organ: colon; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 67.8%; Score 468; DB 13; Length 906;
Best Local Similarity 97.3%; Pred. No. 1.1e-99;
Matches 498; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY 1 CTGCACCTCTCAGTGGCTGACCTCTCTTTGTCATCAGACATCCCTCTCTGGCAGTTGAT 60
Db 300 CTGCACCTCTCAGTGGCTGACCTCTCTTTGTCATCAGACATCCCTCTCTGGCAGTTGAT 359
QY 61 GCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
Db 360 GCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 419
QY 121 GTCACCTCTACAGAGCTTCTATCTCTGGCTTTCATCAGCTGACGCTGACGGTACCTGCC 180
Db 420 GTCACCTCTACAGAGCTTCTATCTCTGGCTTTCATCAGCTGACGCTGACGGTACCTGCC 479
QY 181 ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAAACTCTGGCTGAAAGGAGCTCTAT 240
Db 480 ATTGTCCAGCCACCAACAGTCAAGGCCAAGGAAACTCTGGCTGAAAGGAGCTCTAT 539
QY 241 GTGGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTCACTTTGCGACGTC 300
Db 540 GTGGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTCACTTTGCGACGTC 599
QY 301 AGCCAGGGGACATCAGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 360
Db 600 AGCCAGGGGACATCAGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 659
QY 361 AGCCTGTGATGTGTGTTTCAATTCAGGATATAATAGT-GGGTCTCATCTGCCCGG 419
Db 660 AGCCTGTGATGTGTGTTTCAATTCAGGATATAATAGTGGGGTCTCATCTGCCCGG 719
QY 420 CATGCTCATCTCTCTGTTACTGATCATCATCTCTTAAAGCTGTACACTCCCAAGG-- 476
Db 720 CATGCTCATCTCTCTGTTACTGATCATCATCTCTTAAAGCTGTACACTCCCAAGGGCGC 779
QY 477 CCACCAAGAGCGCAAGGCCCTCAAGACGACAG 508
Db 780 ACCAGAAAGCCAGGCCCTCAAGACGACAG 811

RESULT 11
BX418530
LOCUS
DEFINITION
BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF006VJ16 5-PRIME, mRNA sequence.

ACCESSION
BX418530
VERSION
EST.
SOURCE
Homo sapiens (human)
Homo sapiens

ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF006DE08QPl&cluster=1272.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF006DE08QPl.

FEATURES
source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF006VJ16"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the PCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 55.8%; Score 453.8; DB 13; Length 1201;
Best Local Similarity 84.1%; Pred. No. 2.9e-96;
Matches 573; Conservative 3; Mismatches 85; Indels 20; Gaps 5;
QY 1 CTGCACCTGTGAGTGGTGAACCTCTCTTTGTTCATCAGCTCCCTCTCTGGGAGTTGAT 60
Db 374 CTGCACCTGTGAGTGGTGAACCTCTCTTTGTTCATCAGCTCCCTCTCTGGGAGTTGAT 433
QY 61 GCCATGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACT 120
Db 434 GCCGTGCAAACTGGTACTTTGGGAACTTCTATGCAAGGAGTCCATGTCATCACACA 493
QY 121 GTCAACCTCTACAGCAGCGTTCATCTCTGGCTTTCATCAGCTGACCGGTACCTCGCC 180
Db 494 GTCAACCTCTACAGCAGTGTCTCTCTCTGGCTTTCATCAGTCTGACCGCTACCTGGCC 553
QY 181 ATTGTCCACGCCCAACAGTCAAGGCCAAGGAAACTGTCTGGCTGAAAGGAGTCTAT 240
Db 554 ATCTCCACGCCCAACAGTCAAGGCCAAGGAAAGTGTGGCTGAAAAGGTGGTCTAT 613
QY 241 GTGGGCTCTGATCCAGCCCTCTCTGACTATACCTGACTTTCATCTTTGCGAGTC 300
Db 614 GTGGGCTCTGATCCCTCTCTCTGCTGACTATTCGAGCTTTCATCTTTGCAACG-- 671
QY 301 AGCCAGGGGACATCAGTCAGGGGATGACAGTATCATCTGTGACCGCTTTACCCCGAT 360
Db 672 -----TCAGTRAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAAT 718
QY 361 AGCCTGTGATGTGGTGGTTCATTTCCAGCATATATGTGTGGTCTCATCTCTGCCCGC 420
Db 719 GACTTGTGGTGGTGGTTCAGTTCAGCATCATGTGTGGCTTATCTTGGCTGGT 778
QY 421 ATCTGCTATCTCTCTGTTACTGCTCATCATCTCTAAGCTGTCAACTCCAAAGGGCCAC 480
Db 779 ATTGTCTCTCTCTGCTATTGCTATATCATCTCCAGCTGTCACTCCAAAGGGCCAC 838
QY 481 CAGAGCGCAAGGCCCTCAAGACGACAGTATCTCTATCTCT -AGCTTTCTTTGCTGCTG 539
Db 839 CAGAGCGCAAGGCCCTCAAGACGACAGTATCTCTATCTCTCTCTCTCTCTCTCTCTG 898
QY 540 GCTGCCATATTATGTGGGATCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 599
Db 899 GCTGCCATATTATGTGGGATCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957
QY 600 AGGATGTGATTCGAGAGCATTTGTGCAAGTGGATCTCCATCAGAGGCCCTCGCCTT 659
Db 958 AGGATGTGATTTGAGAACACTGTGCAAGTGGAT-TCCATCAGCGAGGCC--TAGCTT 1014
QY 660 CTTCACCTGTTGCTGAACCC 680
Db 1015 TCTTCMCTGTGTCTGAACCC 1035

RESULT 12
BX462182

QY	421	ATCGTCATCTCTCTGTTACTGCGATCATCTCTAAGCTGTGTCACATCTCAAGGCGCAC	480
Db	762	ATGTGTCATCTCTCTGCTATTGCAATTATCATCTCAAGCTGTGTCACATCTCAAGGCGCAC	821
QY	481	CAGAAAGCGCAAGCGCCTCAAGACGACAGTCATCTCTCATCTCTAGCTTTCTTTGGCTGCTGG	540
Db	822	CAGAAAGCGCAAGCGCCTCAAGACCGACAGTCATCTCTCATCTCTGGCTTTCTTCGCTGTTGG	881
QY	541	CTGCGCATATTATGTGGGATCAGATCGACTCTCTTCATCTCTTTGGGAGTCATCAAGCAA	600
Db	882	CTGCGCTTACTACATTGGGATMA-SATCGACTCTCTTCATCTCTCTGGAATCATMAAGCAA	940
QY	601	GGATGTGACTTCGAGA	616
Db	941	GGGKTGAGCTTTTARA	956

RESULT	13
LOCUS	CD466324
DEFINITION	Leukon2_3 D11.g1.A024 Unstimulated peripheral blood leukocytes N2
ACCESSION	Equus caballus cDNA clone Leukon2_3_D11_A024 5', mRNA sequence.
VERSION	CD466324
KEYWORDS	CD466324.1 GI:31387592
SOURCE	EST.
ORGANISM	Equus caballus (horse)
REFERENCE	Equus caballus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
TITLE	1 (bases 1 to 749)
JOURNAL	Vandenplas,M.; Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M. and Pratt,L.H. An EST database from equine (Equus caballus) unstimulated peripheral blood leukocytes
COMMENT	Unpublished (2003) Other ESTs: Leukon2_3 D11.b1.A024 Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; tissue and RNA were prepared in the Department of Large Animal Medicine, University of Georgia; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA. Seq primer: Sug5 (CTTCGCTCTAAAGCTGCG). Location/Qualifiers 1..749 /organism="Equus caballus" /mol_type="mRNA" /strains="thoroughbred" /db_xref="taxon:9796" /clone="Leukon2_3_D11_A024" /sex="male" /tissue_type="blood" /cell_type="leukocytes" /lab_host="DH10B-T1 phage-resistant E. coli" /clone_lib="Unstimulated peripheral blood leukocytes N2" /note="Organ: circulatory system; Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from unstimulated equine peripheral blood leukocytes isolated from a healthy adult horse. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

FEATURES	source
1..749	

MG1:1077573
Seq primer: -40RP from Gibco
High quality sequence stop: 503.
Location/Qualifiers
1. 582
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:333403"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares mouse NMGB bcell"
/note="Organ: germinal B-cell; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCCTGGTTTTTTTTTTTTTTTTTTT T 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized; constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 63.0%; Score 435; DB 12; Length 582;
Best Local Similarity 100.0%; Pred. No. 5.6e-92;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 CCAGCCCTCCTCCTGACTATACCTGACTTTCATCTTTGGCGAGCTCAGCCAGGGGACATC 315
DB 1 CCAGCCCTCCTCCTGACTATACCTGACTTTCATCTTTGGCGAGCTCAGCCAGGGGACATC 60

QY 316 AGTCAGGGGATGACAGGTACATCTGTGACCGCCCTTACCCCGATAGCTGTGGATGGTG 375
DB 61 AGTCAGGGGATGACAGGTACATCTGTGACCGCCCTTACCCCGATAGCTGTGGATGGTG 120

QY 376 GTGTTTCAATTCCAGCATATATGTTGGTCTCATCTCTGCCCGGATCGTCTCTCTCC 435
DB 121 GTGTTTCAATTCCAGCATATATGTTGGTCTCATCTCTGCCCGGATCGTCTCTCTCC 180

QY 436 TGTTACTGATCATCATCTCTAAGCTGTCACTCCAGGGGCCACAGAGCGCAGGCC 495
DB 181 TGTTACTGATCATCATCTCTAAGCTGTCACTCCAGGGGCCACAGAGCGCAGGCC 240

QY 496 CTCAGACGACAGTCTCATCTCTAGCTTCTTTTGGCTGTGCTGCATATATTATGTG 555
DB 241 CTCAGACGACAGTCTCATCTCTAGCTTCTTTTGGCTGTGCTGCATATATTATGTG 300

QY 556 GGGATCAGCATCGACTCTCTTCATCTCTTTGGGAGTCAATCAAGCAAGGATGTGACTTCGAG 615
DB 301 GGGATCAGCATCGACTCTCTTCATCTCTTTGGGAGTCAATCAAGCAAGGATGTGACTTCGAG 360

QY 616 AGCATTTGTCACAAGTGGATCTCCATCAGAGGGCCCTCGCCTTCTTCCACTGTTCCTTG 675
DB 361 AGCATTTGTCACAAGTGGATCTCCATCAGAGGGCCCTCGCCTTCTTCCACTGTTCCTTG 420

QY 676 AACCCCATCCTCTAT 690
DB 421 AACCCCATCCTCTAT 435

RESULT 15
BX362477 952 bp mRNA linear EST 05-MAY-2003
LOCUS BX362477 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ015YP09 5-PRIME, mRNA sequence.
ACCESSION BX362477
VERSION BX362477.1 GI:30380640
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 952)

ORIGIN

Query Match 63.1%; Score 435.6; DB 14; Length 749;
Best Local Similarity 83.2%; Pred. No. 4.5e-92;
Matches 514; Conservative 0; Mismatches 89; Indels 15; Gaps 1;

QY 73 TGGTACTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACATCTGTCACCTCTAC 132
DB 1 TGGTACTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACATCTGTCACCTCTAC 60

QY 133 AGCAGCGTCTTCATCTCCCTGCGCTTCATCAGCCTGACCGGTACCTCGCCATGTGCCAGCC 192
DB 61 AGCAGTGTCTTCATCTCCCTGCGCTTCATCAGTCTGACCGGTACCTCGGTATCGTCCAGCC 120

QY 193 ACCAACAGTCAAGCCCAAGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
DB 121 ACCAACAGTCAAGCCCAAGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY 253 ATCCAGCGCTTCTTCATCTACCTGACTTTCATCTTTGGCGAGCTCAGCCAGGGGAC 312
DB 181 ATACTGTCTTCTGCTGACTATTTCCCACTTTCATCTTTGCCAAG----- 226

QY 313 ATAGTCAGGGGATGACAGGTACATCTGTGACCGCCCTTTACCCGATAGCTGTGGATG 372
DB 227 -TCAGGAGGGGATGAGGAGGTATATCTGTGACCGCTTTCTACCCGAGTACTGTGGTTG 285

QY 373 GTGGTGTTCATTCACGATATATGTTGGTCTCATCTCTGCCCGGATCGTCTCTCTC 432
DB 286 GTGGTGTTCATTCACGATATATGTTGGTCTCATCTCTGCCCGGATCGTCTCTCTC 345

QY 433 TCCTGTACTGATCATCATCTCTAAGCTGTCACTCCAAAGGCGCCACAGAGCGCAAG 492
DB 346 TCCTGTACTGATCATCATCTCTAAGCTGTCACTCCAAAGGCGCCACAGAGCGCAAG 405

QY 493 GGCCTCAAGACAGTCTCATCTCTAGCTTCTTTTGGCTGTGCTGCTGCTGCTGCTGCTGCT 552
DB 406 GGCCTCAAGACAGTCTCATCTCTAGCTTCTTTTGGCTGTGCTGCTGCTGCTGCTGCTGCT 465

QY 553 GTGGGATCAGCATCGACTCTCTTCATCTCTTTGGGAGTCAATCAAGAGGATGTGACTTC 612
DB 466 ATTTGGATCAGCATCGACTCTCTTCATCTCTCTGGAATCATCCAGGAGGATGTGAGTTT 525

QY 613 GAGACATTTGCAAGTGGTCTCCATCAGAGGGCCCTCGCCTTCTTCCACTGTGCTG 672
DB 526 GAGACATTTGCAAGTGGTCTCCATCAGAGGGCCCTCGCCTTCTTCCACTGTGCTG 585

QY 673 CTGAATCCCATCTCTAT 690
DB 586 CTGAATCCCATCTCTAT 603

RESULT 14
BG145042 582 bp mRNA linear EST 01-FEB-2001
LOCUS BG145042 ut75a01.y1 Soares mouse NMGB bcell Mus musculus cDNA clone
DEFINITION IMAGE:333409 5' similar to SW:CCR4_MOUSE P70658 C-X-C CHEMOKINE
RECEPTOR TYPE 4 ; mRNA sequence.
ACCESSION BG145042
VERSION BG145042.1 GI:12648446
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 582)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 04:56:59 ; Search time 7592.63 Seconds
(without alignments)
10714.982 Million cell updates/sec

Title: US-09-367-052-1
Perfect score: 1877
Sequence: 1 ccatcctaatacagctcact.....aaaaaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1877	100.0	1877	10	D87747 Mus musculus
2	1797.4	95.8	1817	10	BC031665 Mus muscu
3	1683.6	89.7	1809	10	AB000803 Mouse mRN
4	1602.6	85.4	175370	2	AC145861 Pan trogl
5	1602.6	85.4	185215	2	AC147556 Mus muscu
6	1474.8	78.6	1575	10	MMLESTRPT
7	1384.6	73.8	3770	10	MMLESTRGN
8	1216.6	64.8	1223	10	MMU59760
9	1169.4	62.3	190149	2	AC122097
10	1152.6	61.4	1422	10	MMLCR13
11	1143.2	60.9	1180	10	MMLCR12
12	1114	59.4	240931	2	AC107097
13	1055	56.2	3366	10	MMU65580
14	974.4	51.9	1737	6	AR015970
15	974.4	51.9	1737	6	AR060748
16	974.4	51.9	1737	6	AR070433
17	974.4	51.9	1737	6	AR103430
18	974.4	51.9	1737	6	IL3753
19	974.4	51.9	1737	6	IL2455
20	960.4	51.2	1662	9	BC020968
21	958.8	51.1	1362	10	RNU54791
22	952.8	50.8	1667	9	AK129916
23	952.8	50.8	1679	6	AX287116
24	952.8	50.8	1679	6	AX548790
25	952.8	50.8	1679	6	AX705078
26	951.8	50.7	1664	6	E37853
27	951.8	50.7	1664	6	AR222504
28	951.8	50.7	1664	6	AR380668
29	951.8	50.7	1664	9	HUMNM89
30	951.8	50.7	1664	11	G28514
31	947.6	50.5	1645	9	HSNPFYRLA
32	944.8	50.3	1670	6	AX409496
33	944.8	50.3	1670	6	AX663641
34	944.8	50.3	1670	9	HUMGPCR
35	939.8	50.1	1944	6	E33806
36	938.2	50.0	1950	9	AF147204
37	934	49.8	3733	9	HSCXCR4
38	934	49.8	5161	9	AF005058
39	934	49.8	172281	9	AC068492
40	933.6	49.7	1582	6	AX481810
41	932.4	49.7	8747	9	AF052572
42	932.4	49.7	8747	9	HSA224869
43	921.2	49.1	1637	9	HUMSTSR
44	908.8	48.4	1050	10	AF452185
45	907.4	48.3	1689	4	FCA9816

ALIGNMENTS

RESULT 1
D87747
LOCUS D87747 Mus musculus mRNA for murine CXCR-4, complete cds.
DEFINITION D87747
ACCESSION D87747.1 GI:1772445
VERSION D87747.1
KEYWORDS murine CXCR-4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (sites)
AUTHORS Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
Kishimoto,T.

TITLE Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)

MEDLINE 97121456

PUBMED 8962122

REFERENCE 2 (bases 1 to 1877)

AUTHORS Nagasawa, T.

TITLE Molecular cloning of murine PBSF/SDF-1 receptor

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1877)

AUTHORS Nagasawa, T.

TITLE Direct Submission

JOURNAL Submitted (08-SEP-1996) Takashi Nagasawa, Research Institute, Osaka Medical Center, Department of Immunology; Murodocho 840, Izumi, Osaka 590-02, Japan (E-mail: immunol@osk.threewebnet.or.jp, Tel: 0725-56-1220, Fax: 0725-57-3021)

FEATURES

source

1. 1877

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/cell_lines="DW34"

/cell_type="pre-B cell"

/tissue_type="bone marrow"

1. 1877

/gene="murine CXCR-4"

120. .1199

/gene="murine CXCR-4"

/function="PBSF/SDF-1 receptor"

/codon_start=1

/evidence="experimental"

/product="murine CXCR-4"

/protein_id="BAA13451.1"

/db_xref="GI:1772446"

/translation="MEEFISVIVTSNDYSEVGSDYDYNKPCFRDENVHFRFLP
TIFRILGTGNGVLIVMGYQKLRMTDKYRLHLSVADLLFVITLPPWAVDAMA
DWYFGKLCARLIIYVWVYSSVLILAFSLDRYLAIVHATNSQRPKLAEKAVV
GWIPALLITPDIPIADVSQDIDISQDDRYICDRLPVDSLMVWVQFQIHMDLILP
GIVLSVCYIIISKLSHKSHQKELKTTVILILAFACWLPVYVGISDPSILGV
IKQCDRESIVHWISITELAPFHCNLIILAFILGAKPKSAQHALNSMRGSLK
ILSKGKGHSSVSTESSESSPSS"

ORIGIN

Query Match 100.0%; Score 1877; DB 10; Length 1877;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1877; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCTTAATACGACTCACTATAGGCTCGAGCGCGCGCGCGCGAGTGCAGGTAGCAG 60

Db 1 CCATCTTAATACGACTCACTATAGGCTCGAGCGCGCGCGCGAGTGCAGGTAGCAG 60

QY 61 TGACCTCTGAGCGGTTGGTGTCCGGTAACACACCGCTGTAGAGCGAGTGTGCCA 120

Db 61 TGACCTCTGAGCGGTTGGTGTCCGGTAACACACCGCTGTAGAGCGAGTGTGCCA 120

QY 121 TGGAACCGATCAGTGTGAGTATACACTTCTGTATAACTACTCTGAAAGAGTGGGCTG 180

Db 121 TGGAACCGATCAGTGTGAGTATACACTTCTGTATAACTACTCTGAAAGAGTGGGCTG 180

QY 181 GAGACTATGACTCCAAAGAACCCCTGCTCCGGATGAACCGTCCATTTCAATAGGA 240

Db 181 GAGACTATGACTCCAAAGAACCCCTGCTCCGGATGAACCGTCCATTTCAATAGGA 240

QY 241 TCTTCTGCCACCACCATCTACTTTCATCTCTTGACTGGCATAGTCGGCAATGGATTG 300

Db 241 TCTTCTGCCACCACCATCTACTTTCATCTCTTGACTGGCATAGTCGGCAATGGATTG 300

QY 301 TGATCTGGTATGGTATCAGAGAGAGCTAAGGACATGACGGACAAGTACCGGCTGC 360

Db 301 TGATCTGGTATGGTATCAGAGAGAGCTAAGGACATGACGGACAAGTACCGGCTGC 360

QY 361 ACCTGTCACTGGCTGACCTCTCTTTGTCACTCACACTCCCTTCTGGGAGTGTATGCCA 420

Db 361 ACCTGTCACTGGCTGACCTCTCTTTGTCACTCACACTCCCTTCTGGGAGTGTATGCCA 420

QY 421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAAGCTGTCCATATCATCTACACATCA 480

Db 421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAAGCTGTCCATATCATCTACACATCA 480

QY 481 ACCTCTACAGCAGCGTTCTCATCTCTGGCTTTCATCAGCGCTGGACCGGTACCTCGCATTTG 540

Db 481 ACCTCTACAGCAGCGTTCTCATCTCTGGCTTTCATCAGCGCTGGACCGGTACCTCGCATTTG 540

QY 541 TCACGCCACCAACAGCTCAAGGCCCAAGGAATCTCTGGCTGAAAAGGAGTCTATGTGG 600

Db 541 TCACGCCACCAACAGCTCAAGGCCCAAGGAATCTCTGGCTGAAAAGGAGTCTATGTGG 600

QY 601 GGGTCTGATCCAGCCCTCTCTCTGACTATACCTGACTTTCATCTTTGCCAGCTCAGCC 660

Db 601 GGGTCTGATCCAGCCCTCTCTCTGACTATACCTGACTTTCATCTTTGCCAGCTCAGCC 660

QY 661 AGGGGACATCATAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCC 720

Db 661 AGGGGACATCATAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCC 720

QY 721 TGTGGATGGTGGTGTTCATTTCCAGCATATAATGGTGGGTCTCATCTGCCCGCATCG 780

Db 721 TGTGGATGGTGGTGTTCATTTCCAGCATATAATGGTGGGTCTCATCTGCCCGCATCG 780

QY 781 TCATCTCTCTCTGTGACTGACATCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAGA 840

Db 781 TCATCTCTCTCTGTGACTGACATCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAGA 840

QY 841 AGCGCAAGGCCCTCAAGACGACAGTCATCTCTAGCTTCTCTAGCTTCTTGGCTGTGGCTGC 900

Db 841 AGCGCAAGGCCCTCAAGACGACAGTCATCTCTAGCTTCTCTAGCTTCTTGGCTGTGGCTGC 900

QY 901 CATATTATGTGGGATCAGCATCGACTCTCTTCATCTTTGGGAGTCTCAAGCAAGGAT 960

Db 901 CATATTATGTGGGATCAGCATCGACTCTCTTCATCTTTGGGAGTCTCAAGCAAGGAT 960

QY 961 GTGACTTCGAGAGCACTTGTGACAGTGGATCTCCATCAGAGAGCCCTCCCTCTTCTCC 1020

Db 961 GTGACTTCGAGAGCACTTGTGACAGTGGATCTCCATCAGAGAGCCCTCCCTCTTCTCC 1020

QY 1021 ACTGTGCTGAACCCCATCTCTATGCTTCTCTGGGGCCCAAGTTCAAAAGCTCTGCC 1080

Db 1021 ACTGTGCTGAACCCCATCTCTATGCTTCTCTGGGGCCCAAGTTCAAAAGCTCTGCC 1080

QY 1081 AGCATGCACTCACTCCATGACGAGCTCCAGCTCAAGATCTTTCCAAAGGAAAGC 1140

Db 1081 AGCATGCACTCACTCCATGACGAGCTCCAGCTCAAGATCTTTCCAAAGGAAAGC 1140

QY 1141 GGGGTGGACACTCTTCCGTCTCCAGGAGTCAGATCTCCAGTTCCTCAGCTTTCACCTCCAGCTAAC 1200

Db 1141 GGGGTGGACACTCTTCCGTCTCCAGGAGTCAGATCTCCAGTTCCTCAGCTTTCACCTCCAGCTAAC 1200

QY 1201 CCTATGCAAGACTTATATAATATATATATATATATATATATATATATATATATATAT 1260

Db 1201 CCTATGCAAGACTTATATAATATATATATATATATATATATATATATATATATATAT 1260

QY 1261 ACATTTTCCAGATATAAGAGCTGACAGTCTGTACAGTCTTTTCTTTTAAATTTGA 1320

Db 1261 ACATTTTCCAGATATAAGAGCTGACAGTCTGTACAGTCTTTTCTTTTAAATTTGA 1320

QY 1321 CTGTGGGAGTTTATGTTCTCTAGTTTGTGAGGTTTGAATTAATTAATTAATTAATTA 1380

Db 1321 CTGTGGGAGTTTATGTTCTCTAGTTTGTGAGGTTTGAATTAATTAATTAATTAATTA 1380

QY 1381 GTTTTTTCTTTTCTATGTAATGAGCTTACGAGGACCTGTGGCCAGTCTTTAGT 1440

Db 1381 GTTTTTTCTTTTCTATGTAATGAGCTTACGAGGACCTGTGGCCAGTCTTTAGT 1440

QY 1441 AGCTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGGAGAAACTGAACATTCAGAA 1500

[illegible]

Db	781	AGGGCCAC	CAGAGCG	CAAGGCC	CTCAAGAC	GACAGT	CATCCT	CATCTAG	CTTCTTTG	840
Qy	889	CTGTGCTG	CCATATTAT	TGTGGGAT	CAGCAT	CGACTCCT	TCACTC	CTTTTCGG	AGTCA	948
Db	841	CTGTGCTG	CCATATTAT	TGTGGGAT	CAGCAT	CGACTCCT	TCACTC	CTTTTCGG	GGTCA	900
Qy	949	TCAAGCAAG	GATGTGACT	TCGAGAGCA	TTGTGCA	CAAGTGGAT	TCTCCAT	CACAGAG	CCC	1000
Db	901	TCAAGCAAG	GATGTGACT	TCGAGAGCA	TGTGTGCA	CAAGTGGAT	TCTCCAT	CACAGAG	CCCC	960
Qy	1009	TGCGCTTCT	TCAGTGTGCT	CGAACCCAT	CTCTAT	TGCTTCT	CGGGGCCA	AGTTCA	106	
Db	961	TGCGCTTCT	TCAGTGTGCT	CGAACCCAT	CTCTAT	TGCTTCT	CGGGGCCA	AGTTCA	102	

Query Match	89.7%	Score 1683.6	DB 10	Length 1809
Best Local Similarity	97.7%	Pred. No. 0		
Matches 1765	Conservative 0	Mismatches 24	Indels 17	Gaps 5
QY	50	GCAGGTAGCAGTG-ACCCTCTGAGGGGTTTGGTGTCCCGGTAAACCACACCGGCTGTAGAG	108	
DB	1	GCAGGTAGCAGTGAACCTCTGAGGCGTTTGGTGTCCCGGTAAACCACACCGGCTGTAGAG	60	
QY	109	CGAGTGTGTCATGGAAACCGATCAGTGTGAGTATATACATCTTGTGATACTCTGTAAG	168	
DB	61	CGAGTGTGTCATGGAAACCGATCAGTGTGAGTATATACATCTTGTGATACTCTGTAAG	120	
QY	169	AAGTGGGCTCTGAGACTATGATCCCAACAGGAACCTGCTCCGGATGAAAACGTCC	228	
DB	121	AAGTGGGCTCTGAGACTATGATCCCAACAGGAACCTGCTCCGGATGAAAACGTCC	180	
QY	229	ATTTCATAGGATCTTCTCTGCCACCATCTACTTTTCATCATCTTCTTGACTGGCATAGTCG	288	
DB	181	ATTTCATAGGATCTTCTCTGCCACCATCTACTTTTCATCATCTTCTTGACTGGCATAGTCG	240	
QY	289	GCAATGAAATGGTGATCTCTGGTCATGGTTACCAAGAAGACTTAAGSAGATGACGACA	348	
DB	241	GCAATGAAATGGTGATCTCTGGTCATGGTTACCAAGAAGACTTAAGSAGATGACGACA	300	
QY	349	AGTACCGGCTGCACCTGTCAGTGGCTGACCTCTCTTTGTTCATCACACTCCCCCTCTGGG	408	
DB	301	AGTACCGGCTGCACCTGTCAGTGGCTGACCTCTCTTTGTTCATCACACTCCCCCTCTGGG	360	
QY	409	CAGTTGATGGCATGGCTGACGTGTACTTTTGGGAAATTTTGTGTAAAGCTGTCCATATCA	468	
DB	361	CAGTTGATGGCATGGCTGACGTGTACTTTTGGGAAATTTTGTGTAAAGCTGTCCATATCA	420	
QY	469	TCTACACTGTCAACCTCTACAGCAGCTTCTCATCTGGCCCTTCATCAGCCTGGACCGGT	528	
DB	421	TCTACACTGTCAACCTCTACAGCAGCTTCTCATCTGGCCCTTCATCAGCCTGGACCGGT	480	
QY	529	ACCTCGCCATTGTCACGCCACCAACAGTCAAGGCCCAAGAAACTGCTGGCTGAAAAGG	588	
DB	481	ACCTCGCTATTGTCACGCCACCAACAGTCAAGGCCCAAGAAACTGCTGGCTGAAAAGG	540	
QY	589	CAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTTGACTTTCATCTTG	648	
DB	541	CAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTTGACTTTCATCTTG	600	
QY	649	CCGACGTCAAGCAGGGGACATCAGTCAAGGGGATGACAGGTATCATCTGTGACCGCCTTT	708	
DB	601	CCGACGTCAAGCAGGGGACATCAGTCAAGGGGATGACAGGTATCATCTGTGACCGCCTTT	660	
QY	709	ACCCGATAGCTGTGGATGTGGTGTTCATTTCCAGCATATATGTGTGGGTCTCATCC	768	
DB	661	ACCCGATAGCTGTGGATGTGGTGTTCATTTCCAGCATATATGTGTGGGTCTCTGTC	720	
QY	769	TGCCGGGATCGTATCTCTCTGTACTGATCATCATCTCTAAGCTGTACACTCCA	828	
DB	721	TGCCGGGATCGTATCTCTCTGTACTGATCATCATCTCTAAGCTGTACACTCCA	780	
QY	829	AGGGCCACAGAAGCGCAAGGCCCTCAAGACGACAGTCACTCTCACTCACTTCTTTG	888	

RESULT 4

AC145861 175370 bp DNA linear HTG 01-AUG-2003
Pan troglodytes chromosome UNK clone RP43-21D9, *** SEQUENCING IN
PROGRESS ***, 39 unordered pieces.

AC145861
AC145861.1 GI:33386791
HTG; HTGS_PHASE1
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 175370)
Wilson, R.K.
The sequence of Pan troglodytes clone
2 (bases 1 to 175370)
Wilson, R.K.
Direct Submission
Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: C_PF021D09

* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

Contig	Length (bp)	Source
81846	83417	contig of 1572 bp in length
83418	83517	gap of unknown length
83518	83510	contig of 1393 bp in length
84911	85010	gap of unknown length
85011	86788	contig of 1778 bp in length
86789	86888	gap of unknown length
86889	88932	contig of 2004 bp in length
88933	88932	gap of unknown length
88933	90670	contig of 1678 bp in length
90671	90770	gap of unknown length
90771	93207	contig of 2437 bp in length
93208	93307	gap of unknown length
93308	96658	contig of 3351 bp in length
96659	96758	gap of unknown length
96759	100576	contig of 3818 bp in length
100577	100576	gap of unknown length
100677	104988	contig of 4312 bp in length
104989	105088	gap of unknown length
105089	108893	contig of 3805 bp in length
108894	108993	gap of unknown length
108994	112632	contig of 3639 bp in length
112633	112632	gap of unknown length
112733	116328	contig of 3596 bp in length
116329	116328	gap of unknown length
116429	120980	contig of 4552 bp in length
120981	121080	gap of unknown length
121081	126821	contig of 5741 bp in length
126822	126821	gap of unknown length
126922	131331	contig of 4410 bp in length
131332	131331	gap of unknown length
131432	138703	contig of 7274 bp in length
138706	138805	gap of unknown length
138806	145169	contig of 6364 bp in length
145170	145269	gap of unknown length
145270	150696	contig of 5427 bp in length
150697	150796	gap of unknown length
150797	159988	contig of 9192 bp in length
159989	160088	gap of unknown length
160089	167244	contig of 7156 bp in length
167245	167345	gap of unknown length
167345	175370	contig of 8026 bp in length

FEATURES
Location/Qualifiers
1..175370
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="UNK"
/clone="RP43-21D9"
1..5345
misc_feature
1..5345
/note="assembly_name:Contig100"
5446..17195
/note="assembly_name:Contig101"
17296..33250
/note="assembly_name:Contig102"
33351..45646
/note="assembly_name:Contig103"
45747..61992
/note="assembly_name:Contig104"
62093..63423
/note="assembly_name:Contig16"
63524..64666
/note="assembly_name:Contig3"
64767..65994
/note="assembly_name:Contig45"
66095..67430
/note="assembly_name:Contig48"
67531..68982
/note="assembly_name:Contig5"
69083..70891
/note="assembly_name:Contig56"
70992..72172
/note="assembly_name:Contig560"
72273..73614

Db	92497	TTGACTGTTGGGAGTTATGTTTCCCTCTAGTTTGTGAGGTTTGACTTTAAATTTATATAA	92556
Qy	1377	TA-----TTGTTTTTTTGTGTTTTCATGTGAATGAGCGTCTAGCGAGGACCTGTGTGCCAAG	1432
Db	92557	TACTTTTTTTTGTGTTTTCATGTGAATGAGTGTCTAGCGAGGACCTGTGTGCCAAG	92616
Qy	1433	TTCTTAGTAGCTGTTTATCTGTGTGTGAGACGTGTAGAGCTGTAGAGGAGAAACTGAACA	1492
Db	92617	TTCTTAGTAGCTGTTTATCTGTGTGTGAGACGTGTAGAGCTGTAGAGGAGAAACTGAACA	92676
Qy	1493	TTCCAGAACTGTGTGTTAAATTTGAATAAAGCTAGCCGTGATCCTCAGCTGTGTCTCATAA	1552
Db	92677	TTCCAGAACTGTGTGTTAAATTTGAATAAAGCTAGCCGTGATCCTCAGCTGTGTCTCATAA	92736
Qy	1553	TCCTTCATTTCCAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC	1612
Db	92737	TCCTTCATTTCCAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC	92796
Qy	1613	TTGGTTATGCTGTGATGTTGTTGTTGTTGTTGTTT---CTTCTGTGTGTGTTGTTT	1669
Db	92797	TTGGTTATGCTGTGATGTTGTTGTTGTTGTTGTTT---CTTCTGTGTGTGTTGTTT	92856
Qy	1670	TTTTCTGTAAAGATGGCACTTAAACCAAGCCGTAATGTGTGTAGAAATGCTGGGT	1729
Db	92857	TTTTCTGTAAAGATGGCACTTAAACCAAGCCGTAATGTGTGTAGAAATGCTGGGT	92916
Qy	1730	TTTTTTGTTGTTGTTGTTTTCAGTTTTCAGAGCTAGATTGACTTCAGTCCCTACAAAT	1789
Db	92917	TTTTTTGTTGTTGTTTTCAGTTTTCAGAGCTAGATTGACTTCAGTCCCTACAAAT	92976
Qy	1790	GTACAGTCTTTGTTATCATTTGTTTAAATAAAGTCAATGATAAACTTAA	1836
Db	92977	GTACAGTCTTTGTTATCATTTGTTTAAATAAAGTCAATGATAAACTTAA	93023
RESULT 6			
MMLESTRPT		1575 bp	mRNA
LOCUS			linear
DEFINITION			M.musculus mRNA for leukocyte-derived seven transmembrane domain receptor.
ACCESSION			X9582
VERSION			X9582.1
KEYWORDS			lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE			Mus musculus (house mouse)
ORGANISM			Mus musculus
REFERENCE			1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS			Moeppe, B., Frodl, R., Kessler, H. and Gierschik, P.
TITLE			cDNA cloning and genomic organization of a leukocyte-derived seven transmembrane domain receptor (LESTR) from mouse: a murine homologue of the human HIV-1 entry cofactor fusin
JOURNAL			Unpublished
REFERENCE			2 (bases 1 to 1575)
AUTHORS			Moeppe, B.
TITLE			Direct Submission
JOURNAL			Submitted (19-JUL-1996) B. Moeppe, Universitaet Ulm, Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK			Revised by author 11-NOV-96
COMMENT			On Nov 12, 1996 this sequence version replaced gi:1657351.
FEATURES			Location/Qualifiers
source			1..1575
			/organism="Mus musculus"
			/mol_type="mRNA"
			/strain="B6CBA"
			/db_xref="taxon:10090"
			/haplotype="diploid"
			/sex="female"
			/tissue_type="thymus"
			/dev_stage="6-8 weeks old"
			1..1575
			/gene="lestr"
gene			

CDS	90..1169 /gene="lestr" /codon_start=1 /product="leukocyte-derived seven transmembrane domain receptor" /protein_id="CAA67894.1" /db_xref="GI:1666649" /db_xref="GOA:P70658" /xref="SWISS-PROT:P70658" /translation="MEMISYISYSDNYSVEEVSQDYSNKEPCFRDENVHFNRIFLP DWYKFKLCIAVHIIYVNLVSYGVLLAFISLDRLAI VATNSORPKLIAEAAVVA GWIPLLALTDPFI PADVSQGDISQDDRYICDRLYPDSLWMVVFQFQIHVGLVLP GIVLSCCTIILSKLSKHQKXKALKTVIILAPFACWLPYYVGLSIDSPFILLGV IKQCFESIVHKWISITELAFPHCLNLPILYAFILGAKFSSAQHAUNSWSRGSSUK ILSKRGKGGHSSVSSESFSHSS" 1498..1503 /gene="lestr"
	polya_signal
ORIGIN	
	Query Match 78.6%; Score 1474.8; DB 10; Length 1575; Best Local Similarity 98.4%; Pred. No. 5.1e-296; Matches 1525; Conservative 0; Mismatches 12; Indels 13; Gaps 3;
QY	50 GCAGGTAGCAGTGCACCTCTGAGCGGTTTGGTGTCCGGTAAACCAACCGGCTGTAGAGC 109
Db	20 GCAGGTAGCAGTGACCTCTGAGCGGTTTGGTGTCCGGTAAACCAACCGGCTGTAGAGC 79
QY	110 GAGTGTGGCATGGAACGATCAGTGTGAGTATATACATCTTGATACTCTGGAAGA 169
Db	80 GAGTGTGCCATGGAAACGATCAGTGTGAGTATATACATCTTGATACTCTGGAAGA 139
QY	170 AGTGGGGTCTGGAGACTATGACTCTCAACAAGGAACCTGCTCCGGATGAAAACGTCCA 229
Db	140 AGTGGGTTCTGGAGACTATGACTCTCAACAGGAACCTGCTCCGGATGAAAACGTCCA 199
QY	230 TTTCATAGGATCTTCGTGCCCAACATCTACTTCATCATCTTCTTGACTGGCATAGTCG 289
Db	200 TTTCATAGGATCTTCGTGCCCAACATCTACTTCATCATCTTCTTGACTGGCATAGTCG 259
QY	290 CAATCGATTGGTATCGTGGTCATGGGTACCAAGAGAGACTAAGGAGCATGACGGCAA 349
Db	260 CAATGGATTGGTATCGTGGTCATGGGTACCAAGAGAGCTAAGGAGCATGACGGCAA 319
QY	350 GTACCGGCTGCACCTGTGCAGTGGCTGACCTCTCTTTGTTCATCACACTCCCCCTCTGGGC 409
Db	320 GTACCGGCTGCACCTGTGCAGTGGCTGACCTCTCTTTGTTCATCACACTCCCCCTCTGGGC 379
QY	410 AGTTGATGCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCTCATCAT 469
Db	380 AGTTGATGCCATGGCTGACTGGTACTTTGGAAATTTTGTGTAAGGCTGTCTCATCAT 439
QY	470 CTACACTGTCAACCTCTACACAGCGGTTCTCATCTCGGCCCTTCATCAGCGCTGGACCGGTA 529
Db	440 CTACACTGTCAACCTCTACACAGCGGTTCTCATCTCGGCCCTTCATCAGCGCTGGACCGGTA 499
QY	530 CTTGGCATTTGTCAAGCCACCAACAGTCAAAAGCCAGGAACCTGCTGGCTGAAAAGGC 589
Db	500 CTTGGCATTTGTCAAGCCACCAACAGTCAAGGCCAAGAAAACCTGCTGGCTGAAAAGGC 559
QY	590 AGTCTATGTGGCGTCTGGATCCAGCGCCTCTCTGACTATACCTGACTTTCATCTTTGC 649
Db	560 AGTCTATGTGGCGTCTGGATCCAGCGCCTCTCTGACTATACCTGACTTTCATCTTTGC 619
QY	650 CGACGTGACCCAGGGGACATCAGTCTCAGGGGATGACAGGTACATCTGTGACCGGCTTTA 709
Db	620 CGACGTGACCCAGGGGACATCAGTCTCAGGGGATGACAGGTACATCTGTGACCGGCTTTA 679
QY	710 CCCCGATAGCTGTGATGGTGGTTTCATATCCAGCATATATAGTGGGTCTCATCCT 769
Db	680 CCCCGATAGCTGTGATGGTGGTTTCATATCCAGCATATATAGTGGGTCTCATCCT 739
QY	770 GCCCGGCACTGTCATCTCTCTGCTTACTGTCATCATCTCTTAAGCTGTACACTCTCAA 829

[illegible]

REFERENCE
AUTHORS
TITLE
Moepps, B., Prod'l, R., Kessler, H. and Gierschik, P.
cDNA cloning and genomic organization of a leukocyte-derived seven
transmembrane domain receptor (LESTR) from mouse: a murine
homologue of the human Hiv-1 entry cofactor fusin

JOURNAL Unpublished
 REFERENCE 2
 AUTHORS Moeppps,B.
 TITLE Direct Submission
 JOURNAL Submitted (19-JUN-1996) B. Moeppps, Universitaet Ulm,
 Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
 REMARK Revised by [3]
 REFERENCE 3 (bases 1 to 3770)
 AUTHORS Moeppps,B.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-1997) B. Moeppps, Universitaet Ulm,
 Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
 COMMENT On Apr 3, 1997 this sequence version replaced gi:1666646.
 FEATURES
 source
 1..3770
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="B6"
 /db_xref="taxon:10090"
 /haplotype="diploid"
 /tissue_type="thymus"
 /dev_stage="6-8 weeks old"
 41..110
 /number=1
 join(90..110,2306..3364)
 /gene="lestr"
 join(90..110,2306..3364)
 /gene="lestr"
 /codon_start=1
 /product="leukocyte-derived seven transmembrane domain receptor"
 /protein_id="CAA67893.1"
 /db_xref="GI:1666647"
 /db_xref="GOA:P70658"
 /db_xref="SWISS-PROT:P70658"
 /translation="MEPIVSIVYPSDNVSEVSGDYDSNKEPCFRDENVHRIPLP
 TIYFIIFGTGAVHIIYVNLVSSVLLIAPISLDRLYLAIIVHAINSQRPRLKLAERKATV
 DWYFGKFLCKAVHIIYVNLVSSVLLIAPISLDRLYLAIIVHAINSQRPRLKLAERKATV
 GWIPALLLTPTDFADVSQSDISQSDRDYICDLRYPDSLMMVVFQFHIMVGLVLP
 GIVILSCVCIISLKSQKQKALKTIVLILAFACWLPYVIGSIDSTFLLGV
 IKQGDPSIVVHKVTSIPEALAFPHCLNPLVLAFLGAKFKSSAQHALNSMRSGLK
 ILSKRGHSHSVSTESSSSFHSS"
 105..2305
 /misc_feature
 /gene="lestr"
 /note="splice variant/intron"
 111..2305
 /gene="lestr"
 /number=1
 2306..23770
 /number=2
 3693..3698
 /polyA_signal
 ORIGIN
 Query Match 73.8%; Score 1384.6; DB 10; Length 3770;
 Best Local Similarity 97.5%; Pred. No. 3.2e-277;
 Matches 1442; Conservative 0; Mismatches 24; Indels 13; Gaps 3;
 QY 121 TGAACCGATCAGTGTGATATATACATCTTGTGATTAATCTCTGAGAAAGTGGGCTG 180
 DB 2286 TGTCCCTGTTTCTTTCGAGATATACATCTTGTGATTAATCTCTGAGAAAGTGGGCTG 2345
 QY 181 GAGACTATGACTCCAAACAGGAAACCTGCTCCGGGATGAAACGTCCTCAATTTCAATAGGA 240
 DB 2346 GAGACTATGACTCCAAACAGGAAACCTGCTCCGGGATGAAACGTCCTCAATTTCAATAGGA 2405
 QY 241 TCTTCCTGCCACCATCTACTTCATCTCTTTCATCTGCGATAGTCGGCAATGGATTGG 300
 DB 2406 TCTTCCTGCCACCATCTACTTCATCTCTTTCATCTGCGATAGTCGGCAATGGATTGG 2465
 QY 301 TCATCTGGTCTAGGTTTACAGAGAGAGCTAAGGAGCATGACGCAAGTACCGGCTGC 360
 DB 2466 TGATCCTGGTCTAGGTTTACAGAGAGAGCTAAGGAGCATGACGCAAGTACCGGCTGC 2525

361 ACCTGTCACTGGTCACTCCCTCTTTGTTCATCAGACTCCCTCTTGGGAGTTGATGCCA 420
 DB 2526 ACCTGTCACTGGTCACTCCCTCTTTGTTCATCAGACTCCCTCTTGGGAGTTGATGCCA 2585
 QY 421 TGGCTCACTGGTCACTTTGGGAAATTTTGTGTAAGCTCTCCATATCATCTACATGTCA 480
 DB 2586 TGGCTCACTGGTCACTTTGGGAAATTTTGTGTAAGCTCTCCATATCATCTACATGTCA 2645
 QY 481 ACCTCTACAGCAGCTTCTCATCTCCCTGCGCTTCATCAGCTGACCGGTACCTCGCCATTG 540
 DB 2646 ACCTCTACAGCAGCTTCTCATCTCCCTGCGCTTCATCAGCTGACCGGTACCTCGCTATTG 2705
 QY 541 TCCACGCCCAACACAGTCAAGGCCCAAGGAAACTGCTGGCTGAAAGGAGCTTATGTGG 600
 DB 2706 TCCACGCCCAACACAGTCAAGGCCCAAGGAAACTGCTGGCTGAAAGGAGCTTATGTGG 2765
 QY 601 GGCTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTGCGGAGCTCAGCC 660
 DB 2766 GGCTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTGCGGAGCTCAGCC 2825
 QY 661 AGGGGACATCACTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCGATAGCC 720
 DB 2826 AGGGGACATCACTCAGGGGGATGACAGGTACATCTGTGACCGCTTTTACCCGATAGCC 2885
 QY 721 TGTGGATGTGTGTTTCAATTCCAGCATATTAATGGTGGGTCTCATCTTCCCGGATCG 780
 DB 2886 TGTGGATGTGTGTTTCAATTCCAGCATATTAATGGTGGGTCTCGTCTTCCCGGATCG 2945
 QY 781 TCATCTCTCTCTTACTGCTATCATCTCTAAGCTGTCACTCCAAGGGCCACCAGA 840
 DB 2946 TCATCTCTCTCTTACTGCTATCATCTCTAAGCTGTCACTCCAAGGGCCACCAGA 3005
 QY 841 AGCGAAGCGCTCAAGACGAGTCACTCATCTCTAGCTTTCTTTGGCTGTGCTGCTGC 900
 DB 3006 AGCGAAGCGCTCAAGACGAGTCACTCATCTCTAGCTTTCTTTGGCTGTGCTGCTGC 3065
 QY 901 CATATTATGGGGATCAGATCGCTCTCTCATCTTTTGGAGTCATCAAGCAAGGAT 960
 DB 3066 CATATTATGGGGATCAGATCGCTCTCTCATCTTTTGGGGTCATCAAGCAAGGAT 3125
 QY 961 GTGACTTCGAGAGCATTTGCAAGTGGATCTCCATCAAGAGGCTTCTGCTTCTTCC 1020
 DB 3126 GTGACTTCGAGAGCATTTGCAAGTGGATCTCCATCAAGAGGCTTCTGCTTCTTCC 3185
 QY 1021 ACTGTTGCTGAACCCCATCTCTATGCTTCTCGGGGCCAAGTTCAAAGCTCTGCCC 1080
 DB 3186 ACTGTTGCTGAACCCCATCTCTATGCTTCTCGGGGCCAAGTTCAAAGCTCTGCCC 3245
 QY 1081 AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCCCTTCCAAAGGAAAGC 1140
 DB 3246 AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCCCTTCCAAAGGAAAGC 3305
 QY 1141 GGGGTGACACTCTTCCGTCTCCACGAGTCAGATCTCTCCAGTTTTCATCTCAGCTAAC 1200
 DB 3306 GGGGTGACACTCTTCCGTCTCCACGAGTCAGATCTCTCCAGTTTTCATCTCAGCTAAC 3365
 QY 1201 CCTTATGCAAGACTTATATA----ATATATATATATATATATATATATATATATAT 1256
 DB 3366 ACTTATGCAAGACTTAT 3425
 QY 1257 TTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTGTACAG----TTTTTTTTTTTT 1312
 DB 3426 TTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTGTACAGTTTTTTTTTTTTTTT 3485
 QY 1313 TTAATTGACTGTGGAGTTTATGTTCTCTAGTTTCTTGTGAGTTTGAATTAATTTATA 1372
 DB 3486 TTAATTGACTGTGGAGTTTATGTTCTCTAGTTTCTTGTGAGTTTGAATTAATTTATA 3545
 QY 1373 TAAATA-----TTGTTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1427
 DB 3546 TAAATACCTTTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 3605
 QY 1428 CCAAGTCTTAGTAGTGTGTTTATCTGTGTAGGACTGTGAAGACTGTAGAGGAGAAACT 1487

Db 3606 CCAAGTCTTAGTGTGTTTCTGTTGTAGGACTGTAGAACTGTAGAGAAAGAACT 3665
Qy 1488 GAACATTCAGAAATGTGTGTAATTAATAAAGCTAGCGGTGATCCTCAGCTGTGTCTG 1547
Db 3666 GAACATTCAGAAATGTGTGTAATTAATAAAGCTAGCGGTGATCCTCAGCTGTGTCTG 3725
Qy 1548 CATAAATCTTCTTATCCGAGGAGCACCCACCCACCC 1586
Db 3726 CATAAATCTTCTTATCCGAGGAGCACCCACCCACCC 3764

RESULT 8
LOCUS MMU59760 1223 bp mRNA linear ROD 30-SEP-2002
DEFINITION Mus musculus fusin homolog mRNA, complete cds.
ACCESSION U59760
VERSION U59760.1 GI:1527134
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1223)
AUTHORS Reesen,M., Berman,M.A., Benson,J.D., Gerard,C. and Dorf,M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1
co-factor
J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE 97113334
JOURNAL 8955194
PUBMED
REFERENCE 2 (bases 1 to 1223)
AUTHORS Reesen,M., Berman,M.A., Gerard,C. and Dorf,M.E.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA
LOCATION/Qualifiers
FEATURES
source 1..1223
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv"
/db_xref="taxon:10090"
30..1109
/note="seven transmembrane spanning receptor;
chemokine-receptor like"
/codon_start=1
/product="fusin homolog"
/protein_id="AAB07725.1"
/db_xref="GI:1527135"
/translation="MEPISVSIYTSNDYSEVSGDYDSNKEPCFRDENVHFNRIPLP
TIVFIFLTGIVGNGLVILVMGYOKLRSMTDKYRLHLSVADLLPVLTPFWADANA
DMYKFKCNKAVHIITVNLVSSVILAFISLDRLALVHAINSGRERKLLAEKAVY
GWIIFALLUTIPDFIPADVSQDISGDDRYICDRLYDPSLWVVVFQPHIMVGLILP
GVIILSCYIIISKLSHGKQKALKTVILILAFACWLPYVVGISIDSFILLV
IKQGDFFSIVHKWISITEALAFPHCLNPLIYALPLGAKFKSSAQHALNSMRGSLK
ILSKGRGSHSVSESSSPHSS"

Query Match 64.8%; Score 1216.6; DB 10; Length 1223;
Best Local Similarity 99.7%; Pred. No. 2.4e-242;
Matches 1219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 91 ACCACACCGCTGTAGAGCGAGTGTGCCATGGAAACCGATCAGTGTAGTATATACACTT 150
Db 1 ACCACACCGCTGTAGAGCGAGTGTGCCATGGAAACCGATCAGTGTAGTATATACACTT 60

Qy 151 CTGATAACTCTCTGAGAGTGGGGTCTGGAGACTATGACTCCAAAGGAACCCCTGCT 210
Db 61 CTGATAACTCTCTGAGAGTGGGGTCTGGAGACTATGACTCCAAAGGAACCCCTGCT 120

Qy 211 TCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCT 270
Db 121 TCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCT 180

Qy 271 TCTTGACTGGCATAGTCGGCAATGGAATGGTATCCTGGTCAATGGTTACCAAGAGAC 330
Db 181 TCTTGACTGGCATAGTCGGCAATGGAATGGTATCCTGGTCAATGGTTACCAAGAGAC 240
Qy 331 TAAGGAGCATGACGACGACAGTACGGGCTGCACCTGTCAAGTGGCTGACCTCTTGTGTCA 390
Db 241 TAAGGAGCATGACGACGACAGTACGGGCTGCACCTGTCAAGTGGCTGACCTCTTGTGTCA 300
Qy 391 TCACACTCCCTCTTCTGGGCAAGTGTGATGCCATGGCTGATCGTGGTACTTTGGGAAATTTTGT 450
Db 301 TCACACTCCCTCTTCTGGGCAAGTGTGATGCCATGGCTGATCGTGGTACTTTGGGAAATTTTGT 360
Qy 451 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCTGCGCT 510
Db 361 GTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCTGCGCT 420
Qy 511 TCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCAACACAGTCAAGGCCCAAGGA 570
Db 421 TCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCAACACAGTCAAGGCCCAAGGA 480
Qy 571 AACTGTGGCTGAAAGGACAGTCTATGTGGGCTCTGGATCCAGCCCTCTCTCTGACTA 630
Db 481 AACTGTGGCTGAAAGGACAGTCTATGTGGGCTCTGGATCCAGCCCTCTCTCTGACTA 540
Qy 631 TACCTGACTTTCATCTTTGCCGACGTGACGAGGGGGACATCAGTCAGGGGGATGACAGGT 690
Db 541 TACCTGACTTTCATCTTTGCCGACGTGACGAGGGGGACATCAGTCAGGGGGATGACAGGT 600
Qy 691 ACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTCATTTCCAGCAT 750
Db 601 ACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTCATTTCCAGCAT 660
Qy 751 TAATGTGGGTCTCATCTCTGCGCGGATGTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 810
Db 661 TAATGTGGGTCTCATCTCTGCGCGGATGTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Qy 811 CTAAGTGTGACACTCCAGGGGCCACAGAGCGCAAGGCCCTCAAGACGACAGTCAATCC 870
Db 721 CCAAGTGTGACACTCCAGGGGCCACAGAGCGCAAGGCCCTCAAGACGACAGTCAATCC 780
Qy 871 TCATCTCTAGCTTTCTTTGCTGCTGCTGCATATTATGTGGGGATCAGCATCGACTCCT 930
Db 781 TCATCTCTAGCTTTCTTTGCTGCTGCTGCATATTATGTGGGGATCAGCATCGACTCCT 840
Qy 931 TCATCTCTTTGGGAGTCAATCAAGCAAGATGTGACTTCGAGAGCATGTGCAAAAGTGA 990
Db 841 TCATCTCTTTGGGAGTCAATCAAGCAAGATGTGACTTCGAGAGCATGTGCAAAAGTGA 900
Qy 991 TCTCCATCAAGAGGCCCTCGCTTCTTCCACTGTTTGGCTGAACCCCATCTCTATGCT 1050
Db 901 TCTCCATCAAGAGGCCCTCGCTTCTTCCACTGTTTGGCTGAACCCCATCTCTATGCT 960
Qy 1051 TCCTCGGGGCCAAAGTTCAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGT 1110
Db 961 TCCTCGGGGCCAAAGTTCAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGT 1020
Qy 1111 CGAGCTCAAGTCTCTTCCAAAGGAAGGGGTGGAACCTCTTCGGTCTCAACGAGT 1170
Db 1021 CGAGCTCAAGTCTCTTCCAAAGGAAGGGGTGGAACCTCTTCGGTCTCAACGAGT 1080
Qy 1171 CAGAATCTCTCCAGTTCATCTCCAGTAAACCTTATGCAAGAGCTTATATATATATA 1230
Db 1081 CAGAATCTCTCCAGTTCATCTCCAGTAAACCTTATGCAAGAGCTTATATATATATA 1140
Qy 1231 TATATATGATAAGAACTTTTATGTTACATTTCCAGATATAGAGACTGACCACT 1290
Db 1141 TATATATGATAAGAACTTTTATGTTATGTTATCTATTTCCAGATATAGAGACTGACCACT 1200
Qy 1291 CTTGTACAGTTTTTTTTTTTTTTT 1313
Db 1201 CTTGTACAGTTTTTTTTTTTTTTT 1223

TITLE
JOURNAL

DIRECT SUBMISSION

SUBMITTED (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

ON NOV 15, 2002 THIS SEQUENCE VERSION REPLACED GI:23269619.

THE SEQUENCE IN THIS ASSEMBLY IS A COMBINATION OF BAC BASED READS AND WHOLE GENOME SHOTGUN SEQUENCING READS ASSEMBLED USING ATLAS (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). EACH CONTIG DESCRIBED IN THE FEATURE TABLE BELOW REPRESENTS A SCAFFOLD IN THE ATLAS ASSEMBLY (A 'CONTIG-SCAFFOLD'). WITHIN EACH CONTIG-SCAFFOLD, INDIVIDUAL SEQUENCE CONTIGS ARE ORDERED AND ORIENTED, AND SEPARATED BY SIZED GAPS FILLED WITH NS TO THE ESTIMATED SIZE. THE SEQUENCE MAY EXTEND BEYOND THE ENDS OF THE CLONE AND THERE MAY BE WHOLE CONTIGS WITHIN A CONTIG-SCAFFOLD THAT CONSIST ENTIRELY OF WHOLE GENOME SHOTGUN SEQUENCE READS. BOTH END SEQUENCES AND WHOLE GENOME SHOTGUN SEQUENCE ONLY CONTIGS WILL BE INDICATED IN THE FEATURE TABLE.

- - - - - Genomic Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

- - - - - Project Information
Center project name: GKCR
Center clone name: CH230-108A24

- - - - - Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 179135 bases at least Q40
Consensus quality: 181484 bases at least Q30
Consensus quality: 183011 bases at least Q20
Estimated insert size: 187231; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

- - - - -
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

+ + + + + Location/Qualifiers
+ .1..190149
+ /organism="Rattus norvegicus"
+ /mol_type="genomic DNA"
+ /db_xref="taxon:10116"
+ /clone="CH230-108A24"

FEATURES
source

ORIGIN

Query Match 62.3%; Score 1169.4; DB 2; Length 190149;
Best Local Similarity 84.1%; Pred. No. 2.9e-232;
Matches 1465; Conservative 0; Mismatches 156; Indels 122; Gaps 8;

QY 119 CATGGNCCGATCTCATGCCAACAGGAACCCTGCTTCGGGATGAAAACGTCATTTCATATG 178
 |||||
Db 20640 CCGTGTCCCTGTTCTTTTGCGAGATATACACTTCGGATACTTCTCCGAAAGTAGGGGC 20699

QY 179 TGGAGACTATGACTCCTCCAAAGGAACCCTGCTTCGGGATGAAAACGTCATTTCATATG 238
 |||||
Db 20700 TGGAGACTATGACTCCTCCAAAGGAACCCTGCTTCGGGATGAAAACGTAACACAG 20759

QY 239 GATCTTCCTGCCACCATCTACTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATT 298
 |||
Db 20760 GATCTTCCTGCCACCATCTATTTATCATCTTTTATCATCTTTTGATCGCATAGTGGCAATGGTT 20819

Qy	1363	TTAAATTATATAAATATTGTTTTTTT-----GTTGTTTTCATGTGAATGAGCGTC	1411
Db	21895	TTAAATTATATAAATATTGGTTTTTTTTTCTGTGTGTTGTCTGTTTTCAGTGAATGAGTGTC	21944
Qy	1412	TAGGCAGGACCTGTGCGCCAAAGTTCTTACTAGCTGTGTTTATCTGTGTAGGACTCTAGAAC	1471
Db	21945	TAGGCAGGACCTGTGCGCCAAAGTTCTTACTAGCTGTGTTTATCTGTGTAGGACTCTAGAGC	22004
Qy	1472	TGTAGGAGGAAGAACTGTAACATTCGAGAAATGTGTGTTAAATGTAATAAAGCTAGCCGTGA	1531
Db	22005	TGTAGGAGGAAGAACTGTAACATTCGAGAAATGTGTGTTAAATGTAATAAAGCTAGAGTGA	22064
Qy	1532	TGCTCAGCTGTGTCGTGATATCTCTCATTTCCGAGGAGCACCCACCCACCCACCCACCC	1591
Db	22065	CCCTCAGCTGTGTCGTGATATCTCTCTT-----CGCCCG	22098
Qy	1592	CCACCCCATCTTTAAATTTGTTGTTATGCTGTGTGATGTTGTTGTTGTTGTTGTTGTTG	1651
Db	22099	CCACCCCGTTCTGAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	22135
Qy	1652	TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	22139
Db	22136	-----GTTTCTCTGTAGAAGTGCACCTTAAACCAAGGCTGAAATGG	22179
Qy	1712	TGCTAGAAATGCTCGGGTTTTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1771
Db	22180	T-ATAGAATGCTGG-----GTTTCTCTGTTTCAAGAGTAGATTTG	22224
Qy	1772	ACTCAGTCCCTACAAATGTACAGTCTTGTGATATCATTTGTAATAAAGTCAATGATAAA	1831
Db	22225	ATTTCAGTACCTACAAATGTACAGTCTTGTGATATCATTTGTAATAAAGTCAATGATAAA	22284
Qy	1832	CTT 1834	
Db	22285	CTT 2287	
RESULT 10			
LOCUS	MMLCR13		
DEFINITION	Mus musculus lcr-1 gene, exon 2.	DNA	linear
ACCESSION	Z80113	1422 bp	ROD 31-OCT-1996
VERSION	Z80113.1	GI:1655638	
KEYWORDS	CXC chemokine receptor 4; CXCR-4; lcr-1 gene.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Schubel, A., Burgstahler, R. and Lipp, M.		
TITLE	The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN): High expression in thymus and lymphoid tissues		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1422)		
AUTHORS	Lipp, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for Molecular Medicine, MDC, Robert-Rössle-Strasse 10, BERLIN-BUCH, D-13122, GERMANY		
FEATURES			
source	1..1422		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/strain="129SV"		
	/db_xref="taxon:10090"		
	/clone_lib="lambda fix (TM) II"		
	/dev_stage="adult"		
	<1..259		
	/number=1		
intron	260..1422		
gene	/gene="lcr-1"		
exon	260..>1422		
	/gene="lcr-1"		

ORIGIN		/product="CXCR-4"		/standard_name="CXCR-4"		/number=2		/function="receptor for chemokine SDF-1"	
Query Match		61.4%; Score 1152.6; DB 10; Length 1422;							
Best Local Similarity		98.4%; Pred. No. 4.9e-229;							
Matches 1164; Conservative		0; Mismatches 19; Indels 0; Gaps 0;							
QY	121	TGGAACCGATCAGTGTGAGTATATACACTTCTGTGATACTCTCTGTAAGAGAGTGGGTCTG	180						
Db	240	TGTCCTCTGTTTCTTTCAGATATACACTTCTGTGATACTCTCTGTAAGAGAGTGGGTCTG	299						
QY	181	GAGACTATGACTCCAAAGAGAACCTTCTTCGGGATGAAGCTCCATTTCAATAGGA	240						
Db	300	GAGACTATGACTCCAAAGAGAACCTTCTTCGGGATGAAGCTCCATTTCAATAGGA	359						
QY	241	TCCTTCGCGCCACCATCTACTTCAATCTCTTTCGATGCGATAGTCGGCAATGGATTGG	300						
Db	360	TCCTTCGCGCCACCATCTACTTCAATCTCTTTCGATGCGATAGTCGGCAATGGATTGG	419						
QY	301	TGATCTCTGCTCATGGTTTACAGAGAGAGCTTAAGGACATGACGCAAGTACCGGCTGC	360						
Db	420	TGATCTCTGCTCATGGTTTACAGAGAGAGCTTAAGGACATGACGCAAGTACCGGCTGC	479						
QY	361	ACCTGTCTGAGTGGTACCTCTCTTGTCTATCATCACTCCCTCTCTGGSCAGTTGATGCCA	420						
Db	480	ACCTGTCTGAGTGGTACCTCTCTTGTCTATCATCACTCCCTCTCTGGSCAGTTGATGCCA	539						
QY	421	TGGCTGACTGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACTGTCA	480						
Db	540	TGGCTGACTGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCATCTACACTGTCA	599						
QY	481	ACCTCTACAGCAGGTTCTCATCTGCGCTTCATCAGCTGAGCGGTACTCTGCGCAATTG	540						
Db	600	ACCTCTACAGCAGGTTCTCATCTGCGCTTCATCAGCTGAGCGGTACTCTGCGCAATTG	659						
QY	541	TCCAGGCCAACAGTCAAAAGGCCAAGAACTGTCTGTGCTGTAAGAGAGTCTATGTGG	600						
Db	660	TCCAGGCCAACAGTCAAAAGGCCAAGAACTGTCTGTGCTGTAAGAGAGTCTATGTGG	719						
QY	601	CGCTGTGATCCAGCAGCTCTCTCTGCTACTATACCTGACTTCTTTCGCGAGTGTAGCC	660						
Db	720	CGCTGTGATCCAGCAGCTCTCTCTGCTACTATACCTGACTTCTTTCGCGAGTGTAGCC	779						
QY	661	AGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGATAGCC	720						
Db	780	AGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTTACCCCGATAGCC	839						
QY	721	TGTGATGGTGGTGTTCATTCAGCATATATAGTGGTGGTCTCATCTGCGCGGATCG	780						
Db	840	TGTGATGGTGGTGTTCATTCAGCATATATAGTGGTGGTCTCATCTGCGCGGATCG	899						
QY	781	TCATCTCTCTCTGTACTGTACATCATCTCTTAAGCTGTACACTCCAAAGGCCACCCAGA	840						
Db	900	TCATCTCTCTCTGTACTGTACATCATCTCTTAAGCTGTACACTCCAAAGGCCACCCAGA	959						
QY	841	AGCGAAGGCCCTCAAGACGACATCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900						
Db	960	AGCGAAGGCCCTCAAGACGACATCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1019						
QY	901	CATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	960						
Db	1020	CATATTATGTGGGATCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1079						
QY	961	GTGACTTCGAGAGCATTTGTGACAAAGTGGATCTCCATCAGAGAGGCCCTCGCTTCTTCC	1020						
Db	1080	GTGACTTCGAGAGCATTTGTGACAAAGTGGATCTCCATCAGAGAGGCCCTCGCTTCTTCC	1139						
QY	1021	ACTGTTGCTGAACCCCATCTCTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC	1080						
Db	1140	ACTGTTGCTGAACCCCATCTCTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC	1199						

QY	1081	AGCATGCACTCACTCCATGAGCAGAGGCTCCAGGCTCAAGATCCTTTCCAAAGGAAAGC	1140
Db	1200	AGCATGCACTCACTCCATGAGCAGAGGCTCCAGGCTCAAGATCCTTTCCAAAGGAAAGC	1259
QY	1141	GGGTGGACACTCTTCCGCTCTCCAGGAGTCAGATCTCCAGTTTTCACCTCCAGCTAAC	1200
Db	1260	GGGTGGACACTCTTCCGCTCTCCAGGAGTCAGATCTCCAGTTTTCACCTCCAGCTAAC	1319
QY	1201	CTTATGCAAGACACTTAT	1260
Db	1320	CTTATGCAAGACACTTAT	1379
QY	1261	ACATTTTCCAGATATAAGAGACTGACCACTGCTTGTACAGTTT	1303
Db	1380	ACATTTTCCAGATATAAGAGACTGACCACTGCTTGTACAGTTT	1422
RESULT 11			
MMLCR12			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
Source			
gene			
CDS			
exon			
exon			

ORIGIN									
/number=2									
Query Match 60.9%; Score 1143.2; DB 10; Length 1180;									
Best Local Similarity 99.1%; Pred. No. 4.3e-227; Indels 2; Gaps 1;									
Matches 1160; Conservative 0; Mismatches 8;									
QY	136	TGAGTATATACACTCTCTGATAACTACTCTGAAGAAGTGGGCTCTGAGACTATGACTCCA	195						
DB	11	TCAGTATATACACTCTCTGATAACTACTCTGAAGAAGTGGGCTCTGAGACTATGACTCCA	70						
QY	196	ACAAGAACCTCTCTCGGGATGAAGAGCTGCAATTTCAATAGGATCTCTCCGCCACCA	255						
DB	71	ACAAGAACCTCTCTCGGGATGAAGAGCTGCAATTTCAATAGGATCTCTCCGCCACCA	130						
QY	256	TCTACTTTCATCATCTCTGACATGCTAGTGGCAATGATGGTGTGATCTCTGGTCAATG	315						
DB	131	TCTACTTTCATCATCTCTGACATGCTAGTGGCAATGATGGTGTGATCTCTGGTCAATG	190						
QY	316	GTTACCAAGAAAGCTAAGAGCATGACGGAAGTACCGGTGTGACCTGTGACATGGCTG	375						
DB	191	GTTACCAAGAAAGCTAAGAGCATGACGGAAGTACCGGTGTGACCTGTGACATGGCTG	250						
QY	376	ACTCTCTCTTTGTCATCACACTCCCTCTGCGCAGTTGATGCGATGGCTGACTGTACT	435						
DB	251	ACTCTCTCTTTGTCATCACACTCCCTCTGCGCAGTTGATGCGATGGCTGACTGTACT	310						
QY	436	TTGGGAATTTTGTGTAAGGCTGTCCATATCATCTACATGTGTCACCTCTACAGCAGC	495						
DB	311	TTGGGAATTTTGTGTAAGGCTGTCCATATCATCTACATGTGTCACCTCTACAGCAGC	370						
QY	496	TCTCATCTCTGCTTCATCAGCTGAGCGGTGACCTCGCATTTGCCAGCCACCAACA	555						
DB	371	TCTCATCTCTGCTTCATCAGCTGAGCGGTGACCTCGCATTTGCCAGCCACCAACA	430						
QY	556	GTCAAGGCCAAGAAACTCTGCTGAAAGGAGCTCTATGTGGCGCTGTGATCCGAC	615						
DB	431	GTCAAGGCCAAGAAACTCTGCTGAAAGGAGCTCTATGTGGCGCTGTGATCCGAC	490						
QY	616	CCCTCTCTGATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	675						
DB	491	CCCTCTCTGATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	550						
QY	676	AGGGGATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCTGTGATGGTGGTGT	735						
DB	551	AGGGGATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCTGTGATGGTGGTGT	610						
QY	736	TTCAATTCAGCATATATATGTTGGTGTCTATCTGTCGCGGATCGTCACTCTCTCTGT	795						
DB	611	TTCAATTCAGCATATATATGTTGGTGTCTATCTGTCGCGGATCGTCACTCTCTCTGT	670						
QY	796	ACTGCATCATCTCTTAAGCTGTACACTCCAGGSCCACCAGAGCGCAAGCCCTCA	855						
DB	671	ACTGCATCATCTCTTAAGCTGTACACTCCAGGSCCACCAGAGCGCAAGCCCTCA	730						
QY	856	AGAGCAGCATCATCTCTAGCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	915						
DB	731	AGAGCAGCATCATCTCTAGCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	790						
QY	916	TCAGATCGACTCTCTTATCTCTTTGGAGTATCATCAGCAGAGTATGCTTCGAGCA	975						
DB	791	TCAGATCGACTCTCTTATCTCTTTGGAGTATCATCAGCAGAGTATGCTTCGAGCA	850						
QY	976	TTGTGCAAGTGGATCTTCCATCAGAGGCGCTCGCTCTCTTCCACTGTGCTGAACC	1035						
DB	851	TTGTGCAAGTGGATCTTCCATCAGAGGCGCTCGCTCTCTTCCACTGTGCTGAACC	910						
QY	1036	CCATCTCTATGCTCTCTCTGGGCGCAAGTTCAAAAGCTCTGCCAGCATGCTCAACT	1095						
DB	911	CCATCTCTATGCTCTCTCTGGGCGCAAGTTCAAAAGCTCTGCCAGCATGCTCAACT	970						
QY	1096	CCATGAGCAGAGGCTCCAGCTCAAGTCTCTTTCCTTCCAAAGAAAGCGGGGTGACCTCTT	1155						

Db	971	CCATGAGCAGAGCTCCAGCTCAGATCTTCCAAAGAAAGCGGGTGGACACTCTT	1030
QY	1156	CCGCTCTCCAGGAGTCCAGAACTCTCCAGTTTTCACCTCCAGCTAACCTTATCAAGACT	1215
Db	1031	CCGCTCTCCAGGAGTCCAGAACTCTCCAGTTTTCACCTCCAGCTAACCTTATCAAGACT	1090
QY	1216	TAT	1273
Db	1091	TAT	1150
QY	1274	ATAAGAGACTGACCACTCTTGTACAGTTT	1303
Db	1151	ATAAGAGACTGACCACTCTTGTACAGTTT	1180
RESULT 12			
AC107097/c			
LOCUS	AC107097	240931 bp	DNA linear HTG 13-MAY-2003
DEFINITION	Rattus norvegicus clone CH230-175010, *** SEQUENCING IN PROGRESS ***		
ACCESSION	AC107097		
VERSION	AC107097.5 GI:30580624		
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 240931)		
AUTHORS	Munzy, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalbebechi, V., Ayvagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L., Davila, M.I., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M., Gebrerger, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovat, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseghe, H., Lozada, R.J., Lu, X., Ma, J., Madeshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoilemah, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasernak, S., Paul, H., Perez, A., Perez, L., pfannkuch, C., Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, J., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Snares, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,		

Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Keiras, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

TITLE

Direct Submission

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 240931)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (16-JAN-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 240931)

Rat Genome Sequencing Consortium.

REFERENCE

Submitted (13-MAY-2003)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23115278.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKEX

Center clone name: CH230-175010

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 216793 bases at least Q40

Consensus quality: 219254 bases at least Q30

Consensus quality: 221224 bases at least Q20

Estimated insert size: 228277; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* i 240931: contig of 240931 bp in length.

* Location/Qualifiers

1..240931

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-175010"

1..1249

/note="wgs end extension

clone end:Sp6"

complement(5151..5587)

/note="clone boundary

clone end:Sp6

site:ECORI

end_sequence: BH357332"

complement(232579..233220)

/note="clone boundary

clone end:T7

site:ECORI

end_sequence: BH357330"

235468..237558

/note="wgs end extension

clone end:T7"

239450..240931

/note="wgs end extension

clone end:T7"

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match

Best Local Similarity

Matches 1400; Conservative

59.4%; Score 1114; DB 2; Length 240931;

84.2%; Pred. No. 9.8e-221;

0; Mismatches 140; Indels 122; Gaps 8;

QY 200

237558

QY 260

237498

QY 320

237438

QY 380

237378

QY 440

237318

QY 500

237258

QY 560

237198

QY 620

237138

QY 680

237089

QY 740

237033

QY 800

236973

QY 860

236913

QY 920

236853

QY 980

236793

QY

236734

QY

236674

QY

236614

QY

236554

QY

236494

QY

236434

QY

236374

QY

236314

QY

236254

QY

236194

QY

236134

QY

236074

QY

236014

QY

235954

QY

235894

QY

235834

QY

235774

QY 1040 CCTATAGCTTCCTCGGCGCAAGTTCAAAAGCTCTGCCAGCAGCACTCAATCCAT 1099
Db 236733 CCTCTACGCTTCCTCGGCGCAAAATTCAGAGCTCCGCGCAGCTGCACTCAATCCAT 236674
QY 1100 GAGCAGAGCTCCAGCTCAAGATCTTCCAAAGAAAGCGGCTGACACTCTTCGGT 1159
Db 236673 GAGCAGAGCTCCAGCTCAAGATCTTCCAAAGAAAGCGGCTGACACTCTTCGGT 236614
QY 1160 CTCACGAGTCAAGATCTCCAGTTTTCACCTCCAGTAAACCTTATGCAAGACTTATA 1219
Db 236613 CTCACAGAGTCAAGATCTCCAGTTTTCACCTCCAGTAAACCTTATGCAAGACTTATA 236554
QY 1220 TAAT 1270
Db 236553 TAAT 236494
QY 1271 GATATAAGAGCTGACAGCTCTGTACAGTTTTCACCTTTCACCTTTCACCTTTCAC 1330
Db 236493 GATATAAGAGCTGACAGCTCTGTACAGTTTTCACCTTTCACCTTTCACCTTTCAC 236434
QY 1331 TTTATGTTCT 1379
Db 236433 CTCTAGTTTTCGT 236374
QY 1380 TGTCT 1432
Db 236373 TTTTCT 236314
QY 1433 TTTCTAGTCT 1492
Db 236313 TTTCTAGTCT 236254
QY 1493 TTTCCAGATGT 1552
Db 236253 TTTCCAGATGT 236194
QY 1553 TTTCT 1612
Db 236193 TTTCT 236160
QY 1613 TTTGGTTATCTGT 1672
Db 236159 TTTGGTTATCTGT 236139
QY 1673 TTTCT 1732
Db 236138 TTTCT 236086
QY 1733 TTTTGT 1792
Db 236085 TTTTGT 236034
QY 1793 CAGTCTGTAT 1834
Db 236033 CAGTCTGTAT 235992

RESULT 13
LOCUS MMU65580 3366 bp DNA linear ROD 13-DEC-1996
DEFINITION Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION U65580
VERSION U65580.1 GI:1731650
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3366)
AUTHORS Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1 co-factor

J. Immunol. 157 (12), 5455-5460 (1996)
97113334
PUBMED 8955194
REFERENCE 2 (bases 1 to 3366)
AUTHORS Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source
1. 3366
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1. 3366
/gene="CXCR-4"
Join(1..21,2308..3366)
/gene="CXCR-4"
/notes="SDF-1 chemokine receptor"
/codon_start=1
/product="fusin"
/protein_id="AAC52953.1"
/db_xref="GI:3645915"
1. 21
/gene="CXCR-4"
/number=1
1. 15
/gene="CXCR-4"
/notes="alternate exon 1"
16..2307
/gene="CXCR-4"
/notes="alternate intron 1"
22..2307
/gene="CXCR-4"
/number=1
2308..3366
/gene="CXCR-4"
/number=2

exon
exon
intron
intron
exon

ORIGIN
Query Match 56.2%; Score 1055; DB 10; Length 3366;
Best Local Similarity 98.6%; Pred. No. 1.1e-208;
Matches 1064; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 121 TGGAAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGAAGTGGGCTCG 180
Db 2288 TGTCCCTGTTTCTTTTCAGATATACACTTCTGATAACTACTCTGAGAAGTGGGCTCG 2347
QY 181 GAGACTATGATCCCAACAGAAACCTGCTTCGGGATGAAACGTCCTCAATTCATAGGA 240
Db 2348 GAGACTATGATCCCAACAGAAACCTGCTTCGGGATGAAACGTCCTCAATTCATAGGA 2407
QY 241 TCTTCTGCCACCATCTACTTTCATCATCTTCTTGATGCGATAGTCGGCAATGGATTGG 300
Db 2408 TCTTCTGCCACCATCTACTTTCATCATCTTCTTGATGCGATAGTCGGCAATGGATTGG 2467
QY 301 TGATCCTGCTAGTGGTTACAGAGAGAGCTAAGGAGCATGACGCAAGTACCGGCTGC 360
Db 2468 TGATCCTGCTAGTGGTTACAGAGAGAGCTAAGGAGCATGACGCAAGTACCGGCTGC 2527
QY 361 ACCTGTGAGTGGCTGACCTCTCTCTTGTATCATCACTCCCTTCTGGGCAAGTTGATGCCA 420
Db 2528 ACCTGTGAGTGGCTGACCTCTCTCTTGTATCATCACTCCCTTCTGGGCAAGTTGATGCCA 2587
QY 421 TGGTGTAGTGTACTTTGGGAAATTTTGTGAAGCTGTCCATATCATCTACACTCTCA 480

```
Db 2588 TGGCTGACTGGTACTTTTGGGAAATTTTGTGTGAAGGCTGTCCATATCATCTACACTGTCA 2647
Qy 481 ACCTCTACAGACGCTTCTCATCTCTGGCCCTTCATCAGCGCTGACCGGTACCTCGCCATTG 540
Db 2648 ACCTCTACAGACGCTTCTCATCTCTGGCCCTTCATCAGCGCTGACCGGTACCTCGCCATTG 2707
Qy 541 TCACGCCACCAACACAGTCAAAAGGCCAAAGGAACTGCTGCTGAAAGGACGCTATGTGG 600
Db 2708 TCACGCCACCAACACAGTCAAAAGGCCAAAGGAACTGCTGCTGAAAGGACGCTATGTGG 2767
Qy 601 GGGTCTGGATCCAGCCCTCTCTCTGACATATACCTGACTTCACTTTTGGGACGTCAGCC 660
Db 2768 GGGTCTGGATCCAGCCCTCTCTCTGACATATACCTGACTTCACTTTTGGGACGTCAGCC 2827
Qy 661 AGGGGACATCAGTCAAGGGGATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCC 720
Db 2828 AGGGGACATCAGTCAAGGGGATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCC 2887
Qy 721 TGTGGATGCTGTGTGTTTCAATTCACGACATATATATGCTGCTCATCTCTGGCCGATCG 780
Db 2888 TGTGGATGCTGTGTGTTTCAATTCACGACATATATATGCTGCTCATCTCTGGCCGATCG 2947
Qy 781 TCATCTCTCTGTACTGATCATCATCTTAAGCTGTCACTCCAGGGCCACACAGA 840
Db 2948 TCATCTCTCTGTACTGATCATCATCTTAAGCTGTCACTCCAGGGCCACACAGA 3007
Qy 841 AGCGGAAGCCCTCAAGACGACAGTCACTCATCTCATCTAGCTTTCTTTGCTGCTGCTGC 900
Db 3008 AGCGGAAGCCCTCAAGACGACAGTCACTCATCTCATCTAGCTTTCTTTGCTGCTGCTGC 3067
Qy 901 CATATTATGTGGGATCAGCATCGACTCTTCTCATCTCTTTTGGGAGTCACTCAAGGAT 960
Db 3068 CATATTATGTGGGATCAGCATCGACTCTTCTCATCTCTTTTGGGAGTCACTCAAGGAT 3127
Qy 961 GTGACTTTCAGAGCATGTGTGCAAGTGTGATCTCCATCAGAGGCGCTTCGCTTCTTCC 1020
Db 3128 GTGACTTTCAGAGCATGTGTGCAAGTGTGATCTCCATCAGAGGCGCTTCGCTTCTTCC 3187
Qy 1021 ACTGTTGCTGAAACCCATCTCTATGCTTCTCTGGGSCCAAGTCAAAAGCTCTGCC 1080
Db 3188 ACTGTTGCTGAAACCCATCTCTATGCTTCTCTGGGSCCAAGTCAAAAGCTCTGCC 3247
Qy 1081 AGCATCACTCACTCCATGACGAGGCTCCAGCTCAGATCTTTTCCAAAGGAAAGC 1140
Db 3248 AGCATCACTCACTCCATGACGAGGCTCCAGCTCAGATCTTTTCCAAAGGAAAGC 3307
Qy 1141 GGGGTGGACACTCTTCGGTCTCCAGGAGTCAAGATCTTCCAGTTTCACTCCAGCTAA 1199
Db 3308 GGGGTGGACACTCTTCGGTCTCCAGGAGTCAAGATCTTCCAGTTTCACTCCAGCTAA 3366

RESULT 14
AR015970
LOCUS AR015970 1737 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 3 from patent US 5776457.
ACCESSION AR015970
VERSION AR015970.1 GI:3972247
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1737)
AUTHORS Lee J. and Wood W. I.
TITLE Antibodies to human PF4A receptor and compositions thereof
JOURNAL Patent: US 5776457-A 3 07-JUL-1998;
FEATURES
    Location/Qualifiers
        1..1737
            /organism="unknown"
            /mol_type="unassigned DNA"

ORIGIN
Query Match 51.9%; Score 974.4; DB 6; Length 1737;
Best Local Similarity 76.1%; Pred. No. 5.7e-192;
```

```
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;
Qy 50 GCAGGTAGCAGTACACCTCTGAGCGCTTTGGTGTCTCCGTTAAACACACCGCTGTAGAC 109
Db 20 GCGCGCGCAAGATGACGCGCGAGGCGCTGAGTGTCTCAAGTAGCACCGCATCTGGAGAAC 79
Qy 110 GAGTGTTCATGCAACCGCATCAGTGTGAGTATATACACTTCTGTAACTTACTCTGAAGA 169
Db 80 CAGCGGTATACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTTACACCGGGA 134
Qy 170 AGTGGGCTGTGGAGATGATGATCCACACAGGAACCTGCTCCGGGATGAAACGTTCCA 229
Db 135 AATGGGCTCAGGGGACATGATGCTCCATGAGGAACCTGTTCCGTGAGAAATGCTAA 194
Qy 230 TTTCAATAGGATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTGAATGATGCG 289
Db 195 TTTCAATAAATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTAACTGGCATTTGTGG 254
Qy 290 CAATGGATTGGTATCTGTCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 349
Db 255 CAATGGATTGGTATCTGTCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 314
Qy 350 GTACCGGCTGACCTGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 409
Db 315 GTACCGGCTGACCTGTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 374
Qy 410 AGTTGATGCGCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 469
Db 375 AGTTGATGCGCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 434
Qy 470 CTACACGTCAACCTCTACAGCAGCTGTCTCATCTGCGCTTTCATCAGCCTGACCGGTA 529
Db 435 CTACACGTCAACCTCTACAGCAGCTGTCTCATCTGCGCTTTCATCAGCCTGACCGGTA 494
Qy 530 CCTCGCATTTGTCCAGCCACCAACAGTCAAGGCCAAAGGAACTGCTGGCTGAAAGGC 589
Db 495 CCTCGCATTTGTCCAGCCACCAACAGTCAAGGCCAAAGGAACTGCTGGCTGAAAGGC 554
Qy 590 AGTCTATGTGGGCTGTGATCCAGCGCTTCTCTGACTATACCTGACTTCTTCTTTC 649
Db 555 GGTCTATGTGGGCTGTGATCCAGCGCTTCTCTGACTATACCTGACTTCTTCTTTC 614
Qy 650 CGAGCTCAGCCAGGGGACATCAGTCAAGGGGATGACAGGTATCATCTGTGACCGCTTTA 709
Db 615 CAAGC-----TCAGTAGGCGAGATGACAGATATATCTGTGACCGCTTCTA 659
Qy 710 CCCGATAGCTGTGATGAGTGTGATTCATTTCCAGCATATATATGAGTGTGAGTGTGAGT 769
Db 660 CCCCAATGACTTGTGGGCTGTGATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 719
Qy 770 GCGCGCATGCTCATCTCTGTTACTGCATCATCTCTAAAGCTGTCAACTCCAA 829
Db 720 GCGTGGTATTTGTCATCTCTGTTCTGCTGATTTGATTTTCACTTCCAGCTGTCACTCCAA 779
Qy 830 GGGCCACGAGAGCGCAAGCCCTCAAGACGACGATCATCTCATCTTCTTCTTCTTTC 889
Db 780 GGGCCACGAGAGCGCAAGCCCTCAAGACGACGATCATCTCATCTTCTTCTTCTTTC 839
Qy 890 CTGCTGCTGCCATATTTATGTGGGATCAGATCAGTCTCTTCTTCTTCTTCTTCTTCTT 949
Db 840 CTGTTGGCTGCTTACTTACATTTGGGATCAGATCAGTCTCTTCTTCTTCTTCTTCTTCT 899
Qy 950 CAAGCAAGGATGTGACTTTCGAGAGCATTTGTCACAAAGTGGATCTTCCATCAGAGGCCCT 1009
Db 900 CAAGCAAGGATGTGAGTTTGAAGACACTGTGTGCAAGTGGATTTCCATCAGAGGCCCT 959
Qy 1010 GCGCTTCTTCCATCTGTTGCTGAAACCCCATCTCTATGCTTCTTCTGCGGCCAGTTCAA 1069
Db 960 AGCTTTCTTCCATCTGTTGCTGAAACCCCATCTCTATGCTTCTTCTTGGAGCCAAATTA 1019
Qy 1070 AAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCTTCTT 1129
Db 1020 AACTCTGCCAGCAGCATCTCACCTCTGTGAGCAGAGGCTCCAGCCTCAAGATCTTCTT 1079
```

	Qy	1130	CAAAGGAACGGGGTGGACATCTTCCTCGCTCCACGGAGTCAGAATCCTCAGTTTTCA	1189
	Db	1080	CAAAGGAACGGAGGTGGACATTCATCTGTTTCCACTGAGTCTTGAGTTTTCA	1139
	Qy	1190	CTCCAGTCAACCCTTAATGCCAAGACATATAATAATATATATATATATATGATAAAGAACTT	1249
	Db	1140	CTCCAGCTAA-----CACAGATGAAAGACTTTTTTTATACGATAAATACTTTT	1191
	Qy	1250	TTTTATGTTACACATTTTCAGATATAAGAGACTGACCAGTCTTGTACAGTTTTTTTTTTT	1309
	Db	1192	TTTTAAAGTTACACATTTTCAGATATAAAAAGACTGCCAATATTTGTACAGTTTTTATGC	1251
	Qy	1310	TTTTTAATTCACACTGTTGGGAGTTATGTTCTCTAGTTTTTTGTAGGTTTGACTTAATT	1369
	Db	1252	TTCGTTGGAATTTTGT-----CTTGTGTTTCCTTAGTTTGTGGAAGTTTAAITGACTTA	1305
	Qy	1370	ATATAAATATGTTTTTTTTGTTTTCATGTGAATGAGCGTCTAGGCAGACCTGTGGCC	1429
	Db	1306	TTTATA----TAAATTTTTTTTGTTCATATGATGTGTCTAGGCAGACCTGTGGCC	1361
	Qy	1430	AAGTTCCTAGTACGTGTTTATCTGTGTGTTAGGACTGTAGAACCTGTAGAGGAAGAACTGA	1489
	Db	1362	AAGTTCCTAGTTCGTATGTCGTGCTGCTGTTAGGACTGTAGAA-----AAGGGAAGCTGA	1413
	Qy	1490	ACATTCAGAAATGTGTGGTAAATGTAATAAGCTAGCCGTGATCCTCAGCTGTTGTGCA	1549
	Db	1414	ACATTCAGAGCGTGTAGTGAATCAGCTAAAGCTAGAAAAATGATCCCAGCTGTTTATGCA	1473
	Qy	1550	TAATCTCTTCATCCGAGGAGCACCCACCCACCCACCCACCCACCTCTTAAAT	1609
	Db	1474	TAGATAATCTCTCCA-----1488	
	Qy	1610	TGTTTGGTTATGCTGTGNGTAGTGTGTTTGTGTTTTTTTTTGTGTGTTGTTTTT	1669
	Db	1489	-----TTCCCGTGGAAACGTTTTTTCCTGTGTTCTTAAGACGTGAT	1525
	Qy	1670	TTTTCTGTAAGAATGGCACCTTAAACCACAAGCCTGAAATGGTGTGTAAGAAATGCTGGGT	1729
	Db	1526	TTTGCTGTAGAAGATGCACCTTATACCAAGCCCCAAGTGGT-ATAGAANTGCTGG---	1581
	Qy	1730	TTTTTTTTGTTGTTTTTTTTTTTCAGTTTTCAAGAGTAGATGTGACTTCAGTCCCTACAAAT	1789
	Db	1582	-----TTTTTTCAGTTTTTCAGGAGTGGGTGTTTTCAGCACCTAC-AGT	1623
	Qy	1790	GTCAGTCTTGTTATTACATTTGTTAATAAAGTCAATGATAAATTTAAAAAANAANAANA	1849
	Db	1624	GTCAGTCTTGTTATTAAAGTTGTTTAATAAAGTACATGTTTAACTTAAAAAANAANAANA	1683
	Qy	1850	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1877
	Db	1684	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1711

```

RESULT 15
AR060748
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1 (bases 1 to 1737)
Chuntharpai, A., Lee, J., Hebert, C. and Kim, K.Jin.
Antibodies to a human pf4 superfamily receptor
Patent: US 5840856-A 3 24-NOV-1998;
Location/Qualifiers
1. .1737
/organism="unknown"
/mol_type="unassigned DNA"
linear
1737 bp
DNA
Sequence 3 from patent US 5840856.
AR060748
AR060748.1
GI:5987198
Unknown.
Unknown.
Unclassified.

```

Query Match	51.9%;	Score 974.4;	DB 6;	Length 1737;
Best Local Similarity	76.1%;	Pred. No. 5.7e-192;		
Matches 1391;	Conservative	0;	Mismatches 301;	Indels 136; Gaps 10;
QY	50	GCAGGTAGCAGTGCACCCCTCTGAGGCGTTTGGTGTCCGGTACACCACCAACCGGTGTAGAC	109	
DB	20	GGCGGGCGCAAGTGCACGCCGAGGCGCTGAGTGTCCAGTACCCACCGCATCTGGAGAAC	79	
QY	110	GAGTGTGCCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA	169	
DB	80	CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATTAATACACCGAGGA	134	
QY	170	AGTGGGGTCTGGAGACTATGACTCCAAACAAAGGAACCCCTCTTCCGGGATGAAGAACGTCCA	229	
DB	135	AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTTCCGTGAAGAAATGCTAA	194	
QY	230	TTTCAATAGGATTTCCCTGCCACACCATCTACTTCATCTCTTCTTGACTGGGATGTCGG	289	
DB	195	TTTCAATAAAATCTTCTGCCACACCATCTACTTCATCTCTTCTTAACTGGCATTTGTGGG	254	
QY	290	CAATGGATTGGTGATCTCTGGTTCATGGGTTACCAGAAGAAAGCTAAGGAGCATGACGGACAA	349	
DB	255	CAATGGATTGGTTCATCTCGTTCATGGGTTACCAGAAGAAATGAGAAGCATGACGGACAA	314	
QY	350	GTACGGCTGCACCTGTGAGTGGCTGACCTCCTCTTGTGCATCACACTCCCTCTCTGGGC	409	
DB	315	GTACAGGCTGCACCTGTGAGTGGCGACCTCTCTTTGTTCATCACGCTTCCCTTCTGGGC	374	
QY	410	AGTTGATGCCATGGCTGACTGTGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCAT	469	
DB	375	AGTTGATGCCGTGGCAAACTGGTACTTTTGGGAATCTCTTATGCAAGGCAGTCCATGTCTAT	434	
QY	470	CTACACTGTCAACCTCTCAGCAGCGGTTCTCATCTCTGGCTTCATCAGCGCTGACCGGTA	529	
DB	435	CTACACAGTCAACCTCTCAGCAGAGTGCTCTCATCTCTGGCTTCATCAGTCTGACCGCTA	494	
QY	530	CCTCGCATTTGCCAGCCCAACACAGTCAAAAGGCCAAGGAACTGCTGGCTGAAAAGGC	589	
DB	495	CCTGGCATTCTGCCAGCCCAACACAGTCAAGGCCCAAGGAACTGTTGGCTGAAAAGGT	554	
QY	590	AGTCTATGTGGGGCTGTGATCCACGCTCTCTCTGACTATACCTGACTTTCATCTTTGC	649	
DB	555	GGTCTATGTGGCGTCTGATCCCTGCCCTCTGCTGACTATTCCGACATTCTTTGC	614	
QY	650	CGACGTCAAGCGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTACCCGCCCTTTA	709	
DB	615	CAACG-----TCAGTCAGGCAGATGACAGATAATCTGTACCGCTTCTA	659	
QY	710	CCCGATAGCTGTGGATGGTGGTGTTCAAATCCAGCATATAATGGTGGGTCTCATCCT	769	
DB	660	CCCAATGACTTTGGGGTGGTGTGTTCCAGTTTCAGCACATCATGGTGGCCTTATCCT	719	
QY	770	GCCCGCATGTCATCCTCTCCGTTTACTGCATCATCTCTAAGCTGTACACTCCAA	829	
DB	720	GCCTGGTATTGTCATCCTGTCTGTCTATTGCATTATCATCTCCAAGCTGTACACTCCAA	779	
QY	830	GGGCGACCGAAGCGCAAGCCCTCAAGACGACAGTCATCCTCATCTACCTTCTTTTGC	889	
DB	780	GGGCCACCGAAGCGCAAGCCCTCAAGACCGACAGTCATCCTCATCTGCTTCTTCGC	839	
QY	890	CTGCTGGCTGCCATATTATGTGGGGATCAGCATCGACTCCTTCATCCTTTTGGAGTTCAT	949	
DB	840	CTGTTGGCTGCCCTTACTACATTGGGATCAGCATCGACTCCTTCATCCTCTGGAATTCAT	899	
QY	950	CAGACGAGGATGACTTCAGAGCATTTGTCACAAGTGGATCTCCATCACAGAGGCGCT	1009	
DB	900	CAAGCAGGGTGTGAGTTTGAGAACCTGTGCAACAGTGGATTTCATCACCGAGGCGCT	959	
QY	1010	CGCCTTCTTCCACTGTTCCTGAAACCCCACTCTATGCTTCTCTGGGGCCCAAGTTCAA	1069	
DB	960	AGCTTCTTCCACTGTGTGTCTGAAACCCCACTCTATGCTTCTTGGAGCCCAATTTAA	1019	

QY 1070 AAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGATCCCTTTC 1129
Db |||||
1020 AACCTCTGCCAGCAGCACTCACTCTGTGAGCAGAGGCTCCAGCCTCAAGATCCCTTTC 1079
QY 1130 CAAGAGAAAGCGGGGTGACACTCTTCGGTCTCCAGGAGTCAGAAATCCTCCAGTTTCA 1189
Db |||||
1080 CAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCA 1139
QY 1190 CTCAGACTAACCCCTATCCAAAGACTTATATAATATATATATATATATATATATAAGAACTT 1249
Db |||||
1140 CTCAGACTAA-----CACAGATGTAAGAGACTTTTATTATACGATAAATAACTTTT 1191
QY 1250 TTTTATGTATACATTTTCCAGATATAAAGAGACTGACCCAGTCTTGTACAGTTTATTTT 1309
Db |||||
1192 TTTTAAAGTTACACATTTTTCAGATATAAAGAGACTGACCAATATTGTACAGTTTATATGC 1251
QY 1310 TTTTAAATGACTGTGGAGTTTATGTTCTCTAGTCTTTTGTGAGGTTTGCACTTAATTT 1369
Db |||||
1252 TTGTTGGATTTTGT-----CTTGTGTTCTTTAGTTTTTGTGAGTTTAAATTGACTTA 1305
QY 1370 ATATAAATATTGTTTTTGTGTTTTCATGTAATGAGCGCTTAGCGAGGACCTGTGGCC 1429
Db |||||
1306 TTTATA-----TAAATTTTTTTTGTTCATATTGATGTGTCTAGGCAGGACCTGTGGCC 1361
QY 1430 AAGTCTTTAGTAGCTGTTTACTGTGTGAGGACTGTAGAACTGTAGAGGAGAACTGA 1489
Db |||||
1362 AAGTCTTTAGTGTGTGTATGCTCTGTTGAGACTGTAGAA-----AAGGGAAGCTGA 1413
QY 1490 ACATTCACAGATGTGTGTAATGAATAAAGCTAGCCGTGATCTCAGCTGTTGCTCCA 1549
Db |||||
1414 ACATTCACAGAGCGTGTAGTGAATCAGTAAAGCTAGAAATGATCCCCAGCTGTTTATGCA 1473
QY 1550 TAACTCTTCAATCCGAGGAGCACCCACCCACCCACCCACCCACCCACCCACCTCTTAAAT 1609
Db |||||
1474 TAGATAATCTCTCCA----- 1488
QY 1610 TGTGTTGTTATGCTGTGTGATGGTTGTTGTTGTTTGTGTTGTTGTTGTTGTTT 1669
Db |||||
1489 -----TTCCCGTGGAGCGTTTTTCTGTTCTTAAGACGTGAT 1525
QY 1670 TTTTCTGTAAGATGGCACTTAAACCAAGCCTGAAATGGTGTAGAAATGCTGGGGT 1729
Db |||||
1526 TTTGCTGTAGAGATGGCACTTATAACCAAGCCCAAGTGT-ATAGAAATGCTGG--- 1581
QY 1730 TTTTGTGTTGTTGTTGTTTTCAGTCTTCAAGAGTAGATTGACTTCAGTCCCTACAAAT 1789
Db |||||
1582 -----TTTTTCAGTTTTTCAGGAGTGGTTGANTTCAGCACCTAC-AGT 1623
QY 1790 GTACAGTCTTGTATTACATTTGTTAATAAAGTCAATGATAAACTTAAATAAAAAAAAAA 1849
Db |||||
1624 GTACAGTCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAAATAAAAAAAAAA 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAA 1877
Db |||||
1684 AAAAAAAAAAAAAAAAAAAAAA 1711

Search completed: May 17, 2004, 12:10:40
Job time : 7607.63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 03:46:58 ; Search time 804.917 Seconds
(without alignments)
9906.453 Million cell updates/sec

Title: US-09-367-052-1
Perfect score: 1877
Sequence: 1 coactctatacagactcact.....aaaaaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1875.4	99.9	1877	2	AAV46370 Nucleic a
2	1758	93.7	1758	2	AAZ27611 Mouse CX
3	974.4	51.9	1737	2	AAQ80521 Human mon
4	974.4	51.9	1737	2	AAQ80521 Human mon
5	969.6	51.7	1737	2	AAQ80521 Human mon
6	952.8	50.8	1679	4	AAI2801 Human neu
7	952.8	50.8	1679	4	AAI2801 Human neu
8	952.8	50.8	1679	6	ABSS3392 DNA encod
9	952.8	50.8	1679	6	ABV78039 Hypoxia-r
10	952.8	50.8	1679	7	ABX08779 Angiogene
11	952.8	50.8	1679	7	ABX74454 Human CDN
12	952.8	50.8	1679	7	ABZ88886 Nucleotid
13	952.8	50.8	1679	9	ABZ42642 Human CX
14	952.8	50.8	1679	9	ADC98645 Human CX
15	952.8	50.8	1711	9	ABZ35348 Human gen
16	952.8	50.8	1711	9	ADB47320 Human CDN
17	951.8	50.7	1664	3	AA61656 Human CX
18	951.8	50.7	1664	3	AAZ40014 CXCR4 cod
19	950.8	50.7	1679	6	ABSS3750 DNA encod
20	944.8	50.3	1670	6	ABK3803 Human CDN
21	944.8	50.3	1670	6	ABN95645 Gene #214
22	944.8	50.3	1670	7	ACC46765 Human COP
23	939.8	50.1	1944	2	AAI15882 cDNA enco

24	936.8	49.9	1588	2	AAZ27610	Human CX
25	934	49.8	5161	3	AAA34774	Human ade
26	934	49.8	5161	3	AAF20896	Human CCR
27	934	49.8	5161	6	AAH99951	CXCR4 gen
28	934	49.8	5161	7	ABZ96590	Human CCR
29	934	49.8	5161	7	ACA64841	Human CX
30	933.6	49.7	1582	6	ABG64793	CXCR4 rec
31	932.4	49.7	8747	9	ADD67548	Human Ly1
32	905.6	48.2	1050	9	ADB58421	Toxicity
33	905.6	48.2	1050	9	ADB53005	Primary r
34	859	45.8	1225	6	ABZ35630	Human gen
35	859	45.8	1225	7	ACA56637	Human sig
36	847	45.1	1224	6	ABZ35563	Human gen
37	834.6	44.5	1317	2	AAV18357	Human RM3
38	834.6	44.5	1317	2	AAQ91726	Human TTM
39	834.6	44.5	1317	6	ABK54255	Human 7 t
40	833	44.4	1317	2	AAQ66179	Seven tra
41	809.6	43.1	1059	3	AAZ38553	Human CX
42	809.6	43.1	1059	6	ABQ74938	Human CX
43	809.6	43.1	1059	6	AAH99952	CXCR4 enc
44	808	43.0	1059	7	ACC72707	Human can
45	808	43.0	1102	7	ACA64750	Chemokine

ALIGNMENTS

RESULT 1
AAV46370
ID AAV46370 standard; cDNA to mRNA; 1877 BP.

AC AAV46370;

DT 20-NOV-1998 (first entry)

DE Nucleic acid encoding a murine CX chemokine receptor.

KW Mouse; CX chemokine receptor; pre-B cell line DW34;

KW CX chemokine pre-B cell stimulatory factor PBSF/SDF-1; HIV infection;

KW screening; inhibitor; AIDS; ds.

OS Mus sp.

FH Key Location/Qualifiers

FT CDS 120..1199

FT /*tag= a

PN WO9835035-A1.

PD 13-AUG-1998.

PF 07-FEB-1997; 97WO-JP0000299.

PR 07-FEB-1997; 97WO-JP0000299.

XX (SHIO) SHIONOGI & CO LTD.

XX Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;

XX Nakajima T, Yoshie O;

XX WPI: 1998-447232/38.

XX P-PSDB; AA64778.

XX Mouse CX chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory factor - is useful for screening of potential HIV infection and AIDS inhibitors.

XX Claim 3; Page 39-42; 76pp; Japanese.

XX The present sequence encodes a murine CX chemokine receptor which binds to the mouse CX chemokine pre-B cell stimulatory factor PBSF/SDF-1. The nucleic acid is isolated from mouse pre-B cell line DW34. The receptor and cells expressing it can be used in the study and mapping of the

CC mechanism of HIV infection and in screening of potential inhibitors of
CC HIV infection and the development of AIDS
XX

SQ Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;
Query Match 99.9%; Score 1875.4; DB 2; Length 1877;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	CCATCCTTAATACGACTCACTATAGGCTCGAGCGGCCCGCGGCGAGGTGAGGTAGCAG	60
DB	1	CCATCCTTAATACGACTCACTATAGGCTCGAGCGGCCCGCGGCGAGGTGAGGTAGCAG	60
QY	61	TGACCCCTCTGAGCGCTTTGGTGTCTCCGGTAACACCAACCGGCTGTAGAGCGAGTGTGCCA	120
DB	61	TGACCCCTCTGAGCGCTTTGGTGTCTCCGGTAACACCAACCGGCTGTAGAGCGAGTGTGCCA	120
QY	121	TGGAACCGATCAGTGTAGTATATACACTTCTGTATACACTCTCTGATACACTCTCTGAAAGTGGGCTG	180
DB	121	TGGAACCGATCAGTGTAGTATATACACTTCTGTATACACTCTCTGATACACTCTCTGAAAGTGGGCTG	180
QY	181	GAGACTATGACTCCAAAGAAACCGCTCTCCGGGATGAAACCGTCCATTTCAATAGGA	240
DB	181	GAGACTATGACTCCAAAGAAACCGCTCTCCGGGATGAAACCGTCCATTTCAATAGGA	240
QY	241	TCTTCTGCCACCATCTACTTCACTATCTTCTGACTGGCATAGTCGGCAATGGATGG	300
DB	241	TCTTCTGCCACCATCTACTTCACTATCTTCTGACTGGCATAGTCGGCAATGGATGG	300
QY	301	TGATCCTCGTATGAGTGTACAGAGAGAGCTAAGAGCATGACGCAAGAGTACCGGCTGC	360
DB	301	TGATCCTCGTATGAGTGTACAGAGAGAGCTAAGAGCATGACGCAAGAGTACCGGCTGC	360
QY	361	ACCTGTGAGTGGCTGACCTCTCTTGTATCACAACCTCCCTCTTGGGCACTTGAATGCCA	420
DB	361	ACCTGTGAGTGGCTGACCTCTCTTGTATCACAACCTCCCTCTTGGGCACTTGAATGCCA	420
QY	421	TGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGCTGTCCATATCATCTACACTGTCA	480
DB	421	TGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGCTGTCCATATCATCTACACTGTCA	480
QY	481	ACCTCTACAGCAGCGTTCTCATCTCTGGCTTCATCAGCCTGGAACCGGTACCTCGCCATTG	540
DB	481	ACCTCTACAGCAGCGTTCTCATCTCTGGCTTCATCAGCCTGGAACCGGTACCTCGCCATTG	540
QY	541	TCCAGCCCAACAGTCAAGGCCAAGAACTCTGGCTGAAAGGCACTCTATGTGG	600
DB	541	TCCAGCCCAACAGTCAAGGCCAAGAACTCTGGCTGAAAGGCACTCTATGTGG	600
QY	601	GGGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCC	660
DB	601	GGGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCC	660
QY	661	AGGGGACATCAGTCAAGGGGATGACAGTACATCTGTGACCGCTTTACCCGATAGCC	720
DB	661	AGGGGACATCAGTCAAGGGGATGACAGTACATCTGTGACCGCTTTACCCGATAGCC	720
QY	721	TGTGGATGTGTGTTCATTTCCAGCATATATGTTGGTGTCTCATCTCTCCCGCATCG	780
DB	721	TGTGGATGTGTGTTCATTTCCAGCATATATGTTGGTGTCTCATCTCTCCCGCATCG	780
QY	781	TCATCCTCTCTCTGTTACTGCATCATCTCTTAAGCTGTACACTCCAAAGGCCACCCAGA	840
DB	781	TCATCCTCTCTCTGTTACTGCATCATCTCTTAAGCTGTACACTCCAAAGGCCACCCAGA	840
QY	841	AGGGCAGGCCCTCAAGACGACGATCATCTCTAGCTTTCTTTGCTGCTGGCTGC	900
DB	841	AGGGCAGGCCCTCAAGACGACGATCATCTCTAGCTTTCTTTGCTGCTGGCTGC	900
QY	901	CATATTATGTGGGATCAGCATCGACTCCTTCTCATCTTTTGGAGTCTCAAGCAAGAT	960
DB	901	CATATTATGTGGGATCAGCATCGACTCCTTCTCATCTTTTGGAGTCTCAAGCAAGAT	960

RESULT 2
AAZ27611 standard; DNA; 1758 BP.
XX AAZ27611;
XX AC
XX DT
XX 16-DEC-1999 (first entry)

QY	961	GTGACTTCGAGAGCATTGTGCAAGATGGATCTCCATCAGAGAGCCCTCGCTTCTTCC	1020
DB	961	GTGACTTCGAGAGCATTGTGCAAGATGGATCTCCATCAGAGAGCCCTCGCTTCTTCC	1020
QY	1021	ACTGTTGCTGAACCCCAATCCTTATGCTTCTCGGGCCCAAGTTCAAAAGCTCTGCCC	1080
DB	1021	ACTGTTGCTGAACCCCAATCCTTATGCTTCTCGGGCCCAAGTTCAAAAGCTCTGCCC	1080
QY	1081	AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCTTCTCAAGGAAAGC	1140
DB	1081	AGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCTTCTCAAGGAAAGC	1140
QY	1141	GGGGTGACACTTCTCCGTCCTCCAGAGTCAGATCTCCAGTCTTCACTCCAGCTAAC	1200
DB	1141	GGGGTGACACTTCTCCGTCCTCCAGAGTCAGATCTTCTCCAGTCTTCACTCCAGCTAAC	1200
QY	1201	CCTTATGCAAGACITAT	1260
DB	1201	CCTTATGCAAGACITAT	1260
QY	1261	ACATTTTCCAGATATATAAGAGACTGACCACTCTTGTACAGTTTTTTTTTTTAAATGA	1320
DB	1261	ACATTTTCCAGATATATAAGAGACTGACCACTCTTGTACAGTTTTTTTTTTTAAATGA	1320
QY	1321	CTGTTGGAGTTTATGTTCTCTAGTTTTTGTGAGTTTGACTTAATTTATATAATATT	1380
DB	1321	CTGTTGGAGTTTATGTTCTCTAGTTTTTGTGAGTTTGACTTAATTTATATAATATT	1380
QY	1381	GTTTTTTGTGTTTCAATGATGAGCTCTAGGAGGACCTGTGGCCAGTTCTTACT	1440
DB	1381	GTTTTTTGTGTTTCAATGATGAGCTCTAGGAGGACCTGTGGCCAGTTCTTACT	1440
QY	1441	AGCTGTTTATCTGTGTGTAGGACTGTAGAACTGTGTAGAGGAAAGCTGAACATTCAGAA	1500
DB	1441	AGCTGTTTATCTGTGTGTAGGACTGTAGAACTGTGTAGAGGAAAGCTGAACATTCAGAA	1500
QY	1501	TGTTGCTGTAATTTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATATCTCTCA	1560
DB	1501	TGTTGCTGTAATTTGAATAAAGCTAGCCGTGATCCTCAGCTGTTGCTGCATATCTCTCA	1560
QY	1561	TTCGAGGAGCAGCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA	1620
DB	1561	TTCGAGGAGCAGCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA	1620
QY	1621	GCCTGCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1680
DB	1621	GCCTGCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1680
QY	1681	AGATGGCACTTAAACCAAAAGCCTGAAATGSGGTAGAAATGCTGGGTTTTTTTTTGT	1740
DB	1681	AGATGGCACTTAAACCAAAAGCCTGAAATGSGGTAGAAATGCTGGGTTTTTTTTTGT	1740
QY	1741	GTGTTGTTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG	1800
DB	1741	GTGTTGTTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTTG	1800
QY	1801	TATTACATTTGTTTAAATAAAGTCAATCAATAAATTTAAAAAATAAATAAATAAATAA	1860
DB	1801	TATTACATTTGTTTAAATAAAGTCAATCAATAAATTTAAAAAATAAATAAATAAATAA	1860
QY	1861	AAAAAAAAAAAAAAAAA 1877	
DB	1861	AAAAAAAAAAAAAAAAA 1877	

DE		Mouse CXCR4 coding sequence.
XX	CXCR4;	mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW	tissue	repairing agent; vascularisation; ss.
XX		
OS	Mus sp.	
XX	W0948528-A1.	
PN		
XX	30-SEP-1999.	
PD		
XX	23-MAR-1999;	99WO-JP001448.
PF		
XX	24-MAR-1998;	98JP-00095448.
PR	(CHUS) CHUGAI SEIYAKU KK.	
XX	(KISH/) KISHIMOTO T.	
PA		
PI	Kishimoto T,	Nagasawa T, Tachibana X;
PX		
WF	WFI; 1999-591042/50.	
DR	P-PSDB; AAY39994.	
DR		
PT	CXCR4-potentiating agents and methods useful for inhibiting	
PT	neovascularization, and treating solid cancers.	
XX		
PS	Disclosure; Page 50-51; 63pp; Japanese.	
CC	This sequence encodes the mouse CXCR4 protein. The invention relates to	
CC	remedies inhibiting neovascularisation, remedies for solid cancer,	
CC	remedies for diseases pathologically caused by neovascularisation and	
CC	tissue repairing agents containing as the active ingredient a substance	
CC	capable of potentiating CXCR4. Based on a finding that vasculareation is	
CC	inhibited in a CXCR4 knockout mouse, it becomes possible to prepare	
CC	remedies inhibiting vasclarisation which contain as the active for solid	
CC	ingredient a substance capable of potentiating CXCR4, remedies for solid	
CC	cancer, remedies for diseases pathologically caused by neovascularisation	
CC	and tissue repairing agents containing as the active ingredient a	
CC	substance capable of potentiating CXCR4. It is also possible to establish	
CC	methods for treatment with the use of these remedies	
XX		
SQ	Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;	
	Query Match	93.7%; Score 1758; DB 2; Length 1758;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1759; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	120 ATGGAACCGATCAGTGTGAGTATATACACTTCTGTATACTACTCTGAGAAGTGGGCTCT	179
Dd	1 ATGGAACCGATCAGTGTGAGTATATACACTTCTGTATACTACTCTGAGAAGTGGGCTCT	60
Qy	180 GGAGACTATGACTGCCAACGAAGAAACCCTTCCTCGGGATGAAAACGTCCTCAATTCAATAGG	239
Dd	61 GGAGACTATGACTGCCAACGAAGAAACCCTTCCTCGGGATGAAAACGTCCTCAATTCAATAGG	120
Qy	240 ATCTTCTGCGCCACCATCTACTTTTCATCTCTTTCTTGACTGCGCATAGTGGCAATGGAATG	299
Dd	121 ATCTTCTGCGCCACCATCTACTTTTCATCTCTTTCTTGACTGCGCATAGTGGCAATGGAATG	180
Qy	300 GTGATCTGTGCATGGTGTACAGAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTG	359
Dd	181 GTGATCTGTGCATGGTGTACAGAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTG	240
Qy	360 CACCTGFCAGTGGCTGACCTCTCTTTGTATCATCTCCCTTCTGGCGAGTTGATGCC	419
Dd	241 CACCTGFCAGTGGCTGACCTCTCTTTGTATCATCTCCCTTCTGGCGAGTTGATGCC	300
Qy	420 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGCCATATCATCTACACTGTC	479
Dd	301 ATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGCCATATCATCTACACTGTC	360
Qy	480 AACCTCTTACAGCAGCGCTTCTCATCTCTGGCTTTCATCAGCCTGGACCGGTACCTCGCCATT	539

XX	US5440021-A.
PN	
XX	08-AUG-1995.
PD	
XX	25-FEB-1994; 94US-00202056.
PP	
XX	29-MAR-1991; 91US-00677211.
PR	
XX	(HEBE/) HERBERT C.
PA	(CHUN/) CHUNTHARAPAI A.
PA	(KIMK/) KIM K J.
PA	(LEEJ/) LEE J.
XX	
PI	Chuntharapai A, Kim KJ, Hebert C, Lee J;
XX	
WI	WPI; 1995-283151/37.
DR	P-PSTB; AAR80757.
DR	
XX	
PT	New antibodies against interleukin 8 type B receptor - used to treat or
PT	prevent inflammation, also for detecting receptor expression and
PT	purificn.
XX	
PS	Example 2; Col 47-50; 63pp; English.
XX	
CC	Antibodies directed against the interleukin-8 receptor B can be used to
CC	treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid
CC	arthritis and particularly inflammatory bowel disease and chronic lung
CC	inflammation. When immobilised, these antibodies may be used to detect
CC	interleukin-8 receptor B expression in cells and tissues and for affinity
CC	purification of interleukin-8 receptor B from cells. This sequence is an
CC	additional chemokine superfamily receptor which was identified by probing
CC	lambda libraries of genomic DNA from a human monocyte-like cell line (L-
CC	60) and human peripheral blood lymphocytes using a large fragment of the
CC	interleukin-8 type A receptor DNA (See AAQ99006). (Updated on 25-MAR-2003
CC	to correct PF field.)
XX	
SQ	Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T; 0 U; 0 Other;
Query Match	51.9%; Score 974.4; DB 2; Length 1737;
Best Local Similarity	76.1%; Pred. No. 2.9e-178;
Matches 1391; Conservative	0; Mismatches 301; Indels 136; Gaps 10;
QY	50 GCAGGTAGCAGTGACCCCTCTGAGGCCGTTTGGTGCTCCGGTAACCACCAACCGGCTGTAGAGC 109
Db	20 GCOCGGCGGAAGTGCAGCGGAGGGCCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79
QY	110 GAGTGTTCGCATGGAAACCGCATGCTGTGAGTATATACATTCTGTGATACTACTCTCTGAAGA 169
Db	80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAAATACTACACCGAGGA 134
QY	170 AGTGGGGTCTGAGAGCTATGACTCCAACAGGAACCCCTGCTTCCGGGATGAAAACGTCCTCA 229
Db	135 AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTTCCGTGTAGAAATGCTAA 194

250		TTTCAATATATATCTTCTGCCCAACAATCTACTCCATCACTCTCTTACTGGCATTTGGG	254
Db			
290	Qy	CAATGGATTGGTGATCCTGTCATGGGTTACCAAGAAAGACTAAGGAGCANGACGGACAA	349
255	Db	CAATGGATTGGTCATCCTGGTCATGGGTTACCAAGAAAGAACTCAGAAGCATGACGGACAA	314
350	Qy	GTACCGGCTGCACCTGTCAGTGGCTGACGCTCCCTCTTGTGCATCACACTCCCTTCTGGGC	409
315	Db	GTACAGGCTGCACCTGTCAGTGGCGCACCTCTCTTTGTGCATCAGGTTCCCTTCTGGGC	374
410	Qy	AGTTGATGCCATGGCTGCATGGTACTTTGGGAAATTTTGTGTAAAGCTGTCCATATCAT	469
375	Db	AGTTGATGCGGTGGCAAACTGGTACTTTTGGGAACTCTTATGCAAGGCAGTCCATGTTCAT	434
470	Qy	CTACACTGTCAACCTCTACAGCAGCGTTCTCATCTCTGGCCCTTCATCAGCCTGGACCGGTA	529

Db 435 CTACACAGTCAACCTCTACAGCAGTGTCTCTCATCTGGCCCTTCATCATCTGACCGCTA 494
QY 530 CTCTGCGCATGTGCGCAGCCACCAACAGTCAMAAGGCCAAGGAACTGCTGGCTGAAAAGGC 589
Db 495 CTGCGCCATCTGTCAGCCACCAACAGTCAGAGGCCAAGGAGCTGTTGGCTGAAAAGGT 554
QY 590 AGTCATATGCGCGCTGTGATCCGACCCCTCTCTGATATACCTGATCTCATCTTTGCG 649
Db 555 GGTCTATGTGCGCTGTGATCCCTGCGCTCTCTGATATATCCGACTTCATCTTTGCG 614
QY 650 CGACGTCACCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTA 709
Db 615 CAAAG-----TCAGTCAGGACATGACAGATATCTGTGACCGCTTCTA 659
QY 710 CCCCAGATGCTGTGATGCTGTGCTTCAATTCAGCAGATATATGCTGGGTCTCATCT 769
Db 660 CCCCAGATGCTGTGCTGTGCTTCCAGTTTCAGCAGATCATGCTGTGGCTTATCT 719
QY 770 GCGCGCATGCTCATCTCTCTGTTACTGTCATCATCTCTAAGCTGTACATCTCAA 829
Db 720 GCTGTGATGTGCTATCTCTCTGCTGCTATGCTATCTCTCAAGCTGTACATCTCAA 779
QY 830 GGGCCACCAAGAGCGAAGCCCTCAAGACGACAGTCATCTCTCATCTAGCTTTCTTTGC 889
Db 780 GGGCCACCAAGAGCGAAGCCCTCAAGACGACAGTCATCTCTCATCTGGCTTTCTTGC 839
QY 890 CTGCTGCTGCGCATATATGCTGGGATCAGCATCGACTCTCTTCATCTTTGGGAGTCAT 949
Db 840 CTGCTGCTGCGCATATATGCTGGGATCAGCATCGACTCTCTTCATCTTTGGGAGTCAT 899
QY 950 CAAAGCAGATGTGATCTCGAGAGCATTTGTCACAGTGGATCTCCATCAGAGAGCCCT 1009
Db 900 CAAAGCAGAGGTGTGAGTTTGTGAGAACACTGTGCACAGTGAATTCATCAACGAGCCCT 959
QY 1010 CGCTTCTTCCATGCTGTGCTGAGCCCATCTCTATGCTCTCTCGGGCCAACTTCAA 1069
Db 960 AGCTTTCTTCCATGCTGTGCTGAGCCCATCTCTATGCTCTCTCGGGCCAACTTCAA 1019
QY 1070 AAGCTCTGCGCAGCATGCTCACTCACTCCATGACAGAGGCTCAGGCTCAAGATCTTTTC 1129
Db 1020 AAGCTCTGCGCAGCATGCTCACTCTCTGAGCAGAGGCTCAGGCTCAAGATCTTCTC 1079
QY 1130 CAAAGGAAGCGGGGTGACACTCTCTCGTCTCCAGGAGTCAGATCTCCAGTTTCA 1189
Db 1080 CAAAGGAAGCGAGGTGACATCTCTGTTTCCACTGAGTCTGAGTCTCTCAAGTTTCA 1139
QY 1190 CTCAGCTAACCTTATGCAAGACTTATATATATATATATATATATATATATATATAT 1249
Db 1140 CTCAGCTAA-----CACAGTGTAAAGACTTTTTTTATACGATAAATACTTTT 1191
QY 1250 TTTTATGTTACATTTTCCAGATATAAGAGACTGACCAAGTCTGTGACAGTTTCTTTT 1309
Db 1192 TTTTAAAGTTACATTTTCCAGATATAAGAGACTGACCAAGTCTGTGACAGTTTCTTT 1251
QY 1310 TTTTAAAGTTACAGTTTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1369
Db 1252 TTTTAAAGTTACAGTTTATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1305
QY 1370 ATATAATATATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1429
Db 1306 TTTATA-----TAAATTTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1361
QY 1430 AAGTTCTTAGTGTGTTTATCTGTGTGTGAGGACTGTAGAACTGTAGAGGAAGAACTGA 1489
Db 1362 AAGTTCTTAGTGTGTTTATCTGTGTGTGAGGACTGTAGAACTGTAGAGGAAGAACTGA 1413
QY 1490 ACATTCAGAGATGTGTTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1549
Db 1414 ACATTCAGAGATGTGTTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1473
QY 1550 TAATCTCTTCAATTCAGAGGAGCAACCCACCCACCCACCCACCCACCCACCCACCC 1609
Db 1474 TAGATAATCTCTCCA-----1488

QY 1610 TGTTCGTTATGCTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 1669
Db 1489 -----TTCCTGTGAAACGTTTTTTCCTGTTCTTAAGACGTGAT 1525
QY 1670 TTTTCTGTAAGATGCGACTTTAAACCAAGCCTGAAATGTTGTTGTTGTTGTTGTTG 1729
Db 1526 TTTTCTGTAAGATGCGACTTTAATACCAAGCCTGAAATGTTGTTGTTGTTGTTGTT 1581
QY 1730 TTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1789
Db 1582 -----TTTTTCAGTTTTCAGGAGTGGGTGTTGTTTTCAGACCTAC-AGT 1623
QY 1790 GTACAGTCTTGTATTACATTTGTTAAATAAGTCAATGATAAACTTAAAAAATAA 1849
Db 1624 GTACAGTCTTGTATTGTTGTTTAAATAAGTCAATGTTAACTTAAAAAATAA 1683
QY 1850 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1877
Db 1684 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1711

RESULT 5

AAQ29506

ID AAQ29506 standard; DNA; 1737 BP.

XX AAQ29506;

XX 25-MAR-2003 (revised)

DT 12-MAR-1993 (first entry)

XX New platelet factor 4 receptor superfamily member PF4AR1.

XX IL-8R; G-protein coupled receptor family; rhodopsin superfamily;

KW pro-inflammatory cytokine; 8rr.20.15; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 91..1149

FT /*tag= a

FT /*product= "PF4AR1"

XX W09217497-A1.

XX 15-OCT-1992.

XX 23-MAR-1992; 92WO-US002317.

XX 29-MAR-1991; 91US-00677211.

XX 19-DEC-1991; 91US-00810782.

XX (GETH) GENENTECH INC.

XX Lee J, Holmes WE, Wood WI;

XX WPI; 1992-366191/44.

XX P-PSDB; AAR27792.

Isolated human platelet factor 4 super-family receptor polypeptide and corresp. antibodies and DNA - useful as diagnostic and screening agents, and for treating inflammation or PF4AR-mediated disorders.

XX Claim 7; Fig 4; 78pp; English.

CC The IL-8 receptor cDNA sequence was isolated (see AAQ29505) and a 874bp sub-fragment of the coding sequence was used as a probe to screen human cell line HL60 and human peripheral blood lymphocyte cDNA libraries. Two new gene sequences were found that are clearly related to the IL-8 receptor. One of these was contained in combined clone 8rr.20.15 and is predicted to encode an amino acid sequence which is 34% identical with both the high and low affinity IL-8 receptors. See also AAQ37107.

(Updated on 25-MAR-2003 to correct FN field.)

```
XX SQ Sequence 1737 BP; 457 A; 412 C; 370 G; 498 T; 0 U; 0 Other;
Query Match 51.7%; Score 969.6; DB 2; Length 1737;
Best Local Similarity 75.9%; Pred. No. 2.4e-177;
Matches 1388; Conservative 0; Mismatches 304; Indels 136; Gaps 10;
QY 50 GCAGGTAGCAGTACCCCTCTGAGCGTTTGGTCTCCGGTAAACACACAGCGCTGTAGAGC 109
DB 20 GCGCGGCGCAAAATGAGCGCCAGCGGCTGAGTCTCAGTAGCACCGCATCTGGAGAAC 79
QY 110 GAGTGTGCCATGGAACCGCATAGTGTGAGTATATACACTTCTGATAACTCTCTGAAGA 169
DB 80 CAGCGGTACCATGGA-----GGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134
QY 170 AGTGGGCTCGGAGACTATGACTCCAAAGGACCTGCTCCGGGATGAAACCTCCA 229
DB 135 AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCTGTTTCGGTGAAGAAATGCTAA 194
QY 230 TTTCAATAGGATCTTCTGCCCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG 289
DB 195 TTTCAATAAATCTTCTGCCCCACCATCTACTTCATCATCTTCTTAATGGCATTTGGG 254
QY 290 CAATGGATTTGATCTCTGTCTGATGTTTACCAAGAACCTAAGGAGCATGACGACAA 349
DB 255 CAATGGATTTGATCTCTGTCTGATGTTTACCAAGAACCTAAGGAGCATGACGACAA 314
QY 350 GTACCGGCTGCACCTGTGACGTGGCTGACCTCTCTTTGTCATACACTCCCTTCTGGGC 409
DB 315 GTACAGGCTGACCTGTGACGTGGCGACCTCTCTTTGTCATCAGCTTCTCTTCTGGGC 374
QY 410 AGTTGATGCCATGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 469
DB 375 AGTTGATGCCATGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 434
QY 470 CTACACTGTGCACTCTACACAGGCTTCTATCTGCGCTTCTATCAGCTGAGCAGGTA 529
DB 435 CTACACAGTCAACTCTACACAGTGTCTCTATCTGCGCTTCTATCAGTCTGACCGCTA 494
QY 530 CCTCGCCATTTGCCAGGCCAACACAGTCAAGGCCAAGGAACTGTGCTGAAAGGC 589
DB 495 CCTGGCCATGCTCCAGGCCAACACAGTCAAGGCCAAGGAACTGTGCTGAAAGGT 554
QY 590 AGTCTATGTGGGCTGTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCATCTTTC 649
DB 555 GGTCTATGTGGGCTGTGGATCCCTGCGCTCTCTCTGACTATCCGACTTCATCTTTC 614
QY 650 CGAGCTCAGCAGGGGACATCAGTCAAGGGGATGACAGGTATCATCTGTGACCGCTTTA 709
DB 615 CAACG-----TCAGTGAGGACAGATGACAGATATATCTGTGACCGCTTCTA 659
QY 710 CCCGATAGCTGTGGATGGTGGTGTTCATATTCAGCATATATATGTTGGGTCTCATCT 769
DB 660 CCCCATGACTTGTGGTGGTGTTCATATTCAGCATATATATGTTGGGTCTCATCT 719
QY 770 GCCGGCATCTGATCTCTCTCTGTTACTGATCATCATCTCTAAGTGTGACACTCCAA 829
DB 720 GCCTGATATGATCTCTCTCTGTTATGATATATATCTCCAGGTGTGACACTCCAA 779
QY 830 GGGCCACCAAGGCAAGGCCCTCAAGACGACAGTATCTCTCATCTCTAGCTTTCTTTC 889
DB 780 GGGCCACCAAGGCAAGGCCCTCAAGACGACAGTATCTCTCATCTCTAGCTTTCTTTC 839
QY 890 CTGCTGGCTGCTATATATATGTTGGGATCAGCATGCTCTCTCTCTTTTGGGAGTCAT 949
DB 840 CTGTTGGCTGCTTACTACTATGTTGGGATCAGCATGCTCTCTCTCTCTCTGGAATCAT 899
QY 950 CAAGCAAGGATGTACTTCGAGAGCATTTGTGCAAGTGGATCTCCATCAGAGAGGCCCT 1009
DB 900 CAAGCAAGGATGTACTTCGAGAGCATTTGTGCAAGTGGATCTCCATCAGAGAGGCCCT 959
QY 1010 GCGCTTCTTCCACTGTTCCCTGAAACCCCATCTCTATGCGCTTCTCTGCGGCGAAGTTCAA 1069
```

```
DB 960 AGCTTTCTTCCACTGTTGTTCTGAAACCCCATCTCTATGCATTTCTTGGAGCCAAATTTAA 1019
QY 1070 AAGCTCTGCCAGCATGCATCTCACTCACTCATGAGCAGAGGCTCCAGCTCTCAAGATCTTTTC 1129
DB 1020 AACCTCTGCCAGCAGCAGCTCACTCTGTGAGCAGAGGCTCCAGCTCTCAAGATCTTCTC 1079
QY 1130 CAAAGGAAGCGGGGTGGACACTCTTCCGTCTCCAGGAGTCAGAACTCTCCAGTTTCA 1189
DB 1080 CAAAGGAAGCGAGGTGGACATCTCATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCA 1139
QY 1190 CTCGAGCTAAACCTTATGCAAGACTTATATAATATATATATATATATATATATATATATAT 1249
DB 1140 CTCGAGCTAA-----CACAGATGTAAAGACTTTTATACGATAAATAACTTTT 1191
QY 1250 TTTTATGTTACACATTTTCCAGATATAAGAGACTGACAGCTCTTGTACAGTTTTTTTT 1309
DB 1192 TTTTAAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGTACAGTTTATTTGC 1251
QY 1310 TTTTAAATGACTGTGGAGTTTATGCTCTCTAGTTTGTGAGGTTTGTGACTTAAATTT 1369
DB 1252 TTTGTTGGATTTTGT-----CTTGTGTTTCTTTAGTTTGTGAGTTTAAATGACTTA 1305
QY 1370 ATATAAATATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 1429
DB 1306 TTTATA-----TAAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 1361
QY 1430 AAGTCTCTTAGTAGTCTGTTTCTGTGTAGGACTGTGTAGAACTGTAGAGGAAGAACTGA 1489
DB 1362 AAGTCTCTTAGTCTGTATGCTCTGTGTGTAGGACTGTAGAA-----AAGGGAACCTGA 1413
QY 1490 ACATTCAGAAATGCTGGTAAATTTGAATTAAGTCTAGCGTGTATCTCAGCTGTTGCTGCA 1549
DB 1414 ACATTCAGAGGCTGTAGTGAATCAGTAAAGCTAAGAAATGATCCCGAGCTGTTTATGCA 1473
QY 1550 TAATCTCTTCATTCGAGGAGCAGCCACCCACCCACCCACCCACCCACCCACCCACCCATCT 1609
DB 1474 TAGATAATCTCTCCA----- 1488
QY 1610 TGTGTTGTTATGCTGTGTGATGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 1669
DB 1489 -----TTCCCGTGGAACTGTTTCTGTTCTTTAAGACGTGAT 1525
QY 1670 TTTTCTGTAAAGATGGCACTTTAAACCAAGCCCTGAAATGGTGTAGAAATGCTGGGT 1729
DB 1526 TTTGCTGTAGAGATGGCACTTTAAACCAAGCCCTGAAATGGTGTAGAAATGCTGGGT 1581
QY 1730 TTTTCTGTGTTGTTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 1789
DB 1582 -----TTTTTCAGTTTTCAGGAGTGGGTTGATTTTCAGCAGCTAC-AGT 1623
QY 1790 GTACAGTCTGTGTTATGATTTGTTTAAAGTCAATGATAAACTTAAACCAAGGAGGAGG 1849
DB 1624 GTACAGTCTGTGTTATGATTTGTTTAAAGTCAATGTTAACTTAAACCAAGGAGGAGG 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
DB 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```

RESULT 6

AAD12801

ID AAD12801 standard; DNA; 1679 BP.

XX

AC AAD12801;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human neuropeptide Y (NPY) Y3 receptor DNA.

XX

KW Human; neuropeptide Y; NPY; bone disease; bone mass; gene therapy;

KW cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease;

KW fracture; extracellular signal-regulated kinase; ERK; osteoporosis;

KW osteopenia; bone metastasis; neurotransmitter; osteogenic;

NPY Y3 receptor; ds.
 XX Homo sapiens.
 XX Location/Qualifiers
 FH Key 89.1147
 FT CDS /*tag= a
 FT /product= "Human neuropeptide Y (NPY) Y3 receptor"
 XX
 XX WO200153477-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 22-JAN-2001; 2001WO-US002040.
 XX
 XX 20-JAN-2000; 2000US-00489872.
 XX
 XX (BAYU) BAYLOR COLLEGE MEDICINE.
 XX (AMLI/) AMLING M.
 XX
 XX Amling M, Karsenty G, Ducey P;
 XX
 XX WPI; 2001-488709/53.
 XX P-PSDB; AA06690.
 XX
 XX Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing
 XX the level of neuropeptide Y activity in blood or cerebrospinal fluid.
 XX
 XX Example 7; Page 84-85; 102pp; English.
 XX
 XX The present invention relates to a method for treatment or prevention of
 XX bone diseases characterized by loss of bone mass, comprises administering
 XX to a mammal a compound that lowers the level of neuropeptide Y (NPY) in
 XX the serum or cerebrospinal fluid (CSF) or a compound that lowers the
 XX level of inositol phosphate (IP) or extracellular signal-regulated kinase
 XX (ERK). The method is specifically used to treat (including by gene
 XX therapy) or prevent osteoporosis, osteopenia or Paget's disease, but may
 XX also be used e.g., in cases of fractures or bone metastases. These
 XX diseases may also be diagnosed by detecting elevated NPY levels.
 XX including monitoring of treatment, assessing efficacy of compounds in
 XX clinical trials and for identifying subjects at risk. The present
 XX sequence is a human NPY Y3 receptor DNA
 XX
 XX Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
 Query Match 50.8%; Score 952.8; DB 4; Length 1679;
 Best Local Similarity 80.2%; Pred. No. 4.1e-174;
 Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;
 44 GCAGGTGCGAGGTAGCAGTACCCCTCTGAGCGGTTTGGTCTCCGGTAACCACCGGCTG 103
 12 GCGGCGACGAGGTAGCAAGTGAACGCGCGGGGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
 104 TAGAGCGAGTGTGCCATGGACCGCATCAGTGTGAGTATATACATCTCTGTAACTACTC 163
 72 GAGAACCGCGGTACCATGGA-----GGGATCAGTATATACATCTCAGATACTACAC 126
 164 TGAAGAGTGGGGTCTGGAGACATATGACTCCCAAGAGAACCTGCTCCGGGATGAAAA 223
 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCTGTTCCGTGGAAGAAA 186
 224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTCATCATCTTCTGACTGGCAT 283
 187 TGTAAATTTCAATAAATCTTCTCGCCACCATCTACTCATCATCTTCTTAACCTGGCAT 246
 284 AGTCGCAATGGATTTGGTATCTGTGATGGTTACCGAAGAGCTAAGGAGCATGAC 343
 247 TGTGGGCAATGGATTTGGTATCTGTGATGGTTACCGAAGAGCTAAGGAGCATGAC 306
 344 GGCACAGTACCGGCTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
 307 GGACAAAGTACAGGCTGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366

QY 404 CTGGCAGTTGATGCTCATGGCTGACTGTACTTTTGGGAAATTTTGTGTGAAGCTGTCCA 463
 DB 367 CTGGCAGTTGATGCTCGGTGGCAAACTGGTACTTTTGGGAACTTCTCTATGCAAGCGAGTCCA 426
 QY 464 TATCATCTACACTGTCAACCTCTTACAGCAGCGTTCTATCTCTGGGCTTTCATCAGCCTGGA 523
 DB 427 TGTCACTACACAGTCAACCTCTTACAGCAGTGTCTCTCACTCTCTGGGCTTTCATCAGTCTGA 486
 QY 524 CCGGTACCTCGCATTTGTCAGCGCACCAACAGTCAAGAGCCCAAGGAACTGCTGGCTCA 583
 DB 487 CCGCTACCTGGCCATCGTCCAGCCCAACAGTCAAGAGCCCAAGGAACTGCTGGCTCA 546
 QY 584 AAAGGCACTTATGTGGGGTCTGGAATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
 DB 547 AAAGTGGTCTATGTGGGGTCTGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
 QY 644 CTTTGGCGAGTCAAGCCAGGGGGACATCAGTCAGGGGATGACAGGTACATCTCTGACCG 703
 DB 607 CTTTGGCAAG-----TCAGTGGGAGATGACAGATATATCTCTGACCG 651
 QY 704 CTTTACCCCGATAGCTGTGGATGGTGTGTCTCAATTCAGCATATATATGTTGGTCT 763
 DB 652 CTTTACCCCAATGACTTGTGGGTGTGTGTCTCAGTTTTCAGCATCATGTTGGCT 711
 QY 764 CATCTGCCCGGATCGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
 DB 712 TATCTGCTGTGTATGTCT 771
 QY 824 CTCGAAGGCGCACGAGCGCAAGCGCTCTCAAGACGACAGTCACTCTCTCTCTCTCTCTCT 883
 DB 772 CTCGAAGGCGCACGAGCGCAAGCGCTCTCAAGACGACAGTCACTCTCTCTCTCTCTCTCT 831
 QY 884 CTTTGGCTGTCTGCCATATTTATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCT 943
 DB 832 CTTTGGCTGTCTGCCATATTTATGTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCT 891
 QY 944 AGTCATCAACAGAGGATGACTTCGAGGACATGTGCAAGTGTGCAAGTGTGCTCTCTCTCT 1003
 DB 892 AATCATCAAGCAAGGGTGTGAGTTTGAGAACATCTGTGCAAGTGTGATTTCTCTCTCTCT 951
 QY 1004 GGGCTCTCGCT 1063
 DB 952 GGGCT 1011
 QY 1064 GTTCAAAAGCTCTGCCAGCATGCACTCACTCACTCACTGAGCAGAGGCTCCAGCCTCAAGAT 1123
 DB 1012 ATTTAAACCTCTGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1071
 QY 1124 CTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCAGCGAGTCAAGATCTCTCCAG 1183
 DB 1072 CTTCTCCAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG 1131
 QY 1184 TTTTCACTCCAGTAACTCTTATGCAAGACTTATATAATATATATATATATATATATATAA 1243
 DB 1132 TTTTCACTCCAGTAACTCAACAGATGTAAGACTTTTT-----TTTATACATATAAATA 1183
 QY 1244 GAACCTTTTATGTATACACATTTTCCAGATATAAGAGACTGACAGCTCTTGTACAGTTTT 1303
 DB 1184 ACTTTTATTTTAAAGTACACATTTTTCAGATATAAGAGACTGACCAATATTTGTACAGTTT 1243
 QY 1304 TTTTATTTTAAATGACTGTGGAGTTTATGTTCTCTAGTTTGTGAGGTTTGTGAGGTTTACT 1363
 DB 1244 TATTGCTGTGTGATTTTGT-----CTTGTGTCTTTTGTAGTTTGTGAGGTTTAAAT 1297
 QY 1364 TAAATTTATATAATATTTTGT 1423
 DB 1298 GACTTATTTATA-----TAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1353
 QY 1424 GTGGCCAAAGTTCTTTAGTGTGTTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1483
 DB 1354 GTGGCCAAAGTTCTTTAGT 1405
 QY 1484 AACTGAACATTCAGAGT 1543

```

Db 1406 AACTGAACTCCAGAGCGTGTAGTGAATCATCGTAAGCTAGAAATGATCCCGAGCTGTT 1465
QY 1544 GCTGCATA 1551
Db 1456 TATGCATA 1473

RESULT 7
AA165467
ID AA165467 standard; DNA; 1679 BP.
XX
AC AA165467;
XX
DT 10-DEC-2001 (first entry)
XX
DE Nucleotide sequence of a human polynucleotide.
XX
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.
XX
OS Homo sapiens.
XX
PN WO200164752-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006322.
XX
PR 02-MAR-2000; 2000US-00517605.
XX
PA (UYNV ) UNIV NEW YORK STATE.
XX
PA (UYNV - ) UNIV NIJMEGEN.
XX
PI Litman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
PI WPI; 2001-602565/68.
XX
XX
XX An antibody for the treatment or prevention of HIV-infection comprises a
PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
PT DC-SIGN due to concomitant conformational change.
XX
PS Disclosure; Page 126-127; 131pp; English.
XX
XX The specification describes an antibody which is specific for an
CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
CC is exposed upon gp120 binding of DC-SIGN due to concomitant
CC conformational change. DC-SIGN is a receptor that is specifically
CC expressed on dendritic cells and facilitates infection of T lymphocytes
CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
CC with high affinity. The antibody of the invention inhibits the trans
CC enhancement of HIV entry into a T cell or macrophage facilitated by
CC dendritic cells. The antibody is useful to treat or prevent HIV
CC infection. The present sequence represents a human polynucleotide, which
CC is used in the course of the invention
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 50.8%; Score 952.8; DB 4; Length 1679;
Best Local Similarity 80.2%; Pred. No. 4.1e-174;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;
QY 44 GCAGGTGCGAGGTAGCAGTACCCCTCTGAGCGCTTTGGTCTCCCGTAACCAACCGGCTG 103
Db 12 GCGGCAGCAGGTAGCAAAAGTACGCGCGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACATCTCTGATACTACTC 163
Db 72 GGAACACCGGTTACCATGGA-----GGGATCAGTATATACATTCAGATACTACAC 126
QY 164 TGAAGAAGTGGGGTCTGGGACATATGACTCCCAAGGAACCCCTGCTCCGGGATGAAAA 223

```

```

Db 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTCCGTGAAGAAAA 186
QY 224 CGTCCATTTCAATAGATCTTCTGCCACCACTACTTACTCATCTCTCTTCTGACCTGGCAT 283
Db 187 TGTAAATTTCAATAAAATCTTCTGCCACCACTACTTCTCATCTCTCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGATGATGATCTCTGCTCATCTGGGTACCAGAAAGAGCTAAGAGCATGAC 343
Db 247 TGTGGCAATGATGATGATCTCTGCTCATCTGGGTACCAGAAAGAACTGAGAAGCATGAC 306
QY 344 GGACAAAGTACCGGCTGCACCTGTCTAGTGGCTGACCTCTCTCTTCTTGTGTCATCACAATCCCTTT 403
Db 307 GGACAAAGTACCGGCTGCACCTGTCTAGTGGCTGACCTCTCTCTTGTGTCATCACGCTTCCCTTT 366
QY 404 CTGGCAGTGTGATGCTCATCTGGCTGACCTGTGCTACTTGTGGAAATTTTGTGTGAAGCTGTCCA 463
Db 367 CTGGCAGTGTGATGCTCATCTGGCTGCAAACTGTGTACTTGTGGAACTTCTTATGCAAGGAGTCCA 426
QY 464 TATCATCTACATGTCAACCTCTACAGCAGCGTTCTCTCTCTGGCCCTTCTCATGACCTGGA 523
Db 427 TGTCACTACAGCTCAACCTCTACAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 486
QY 524 CGGTACTCTCGCATTTGTCACGCCCAACCAAGTCAAGGCCCAAGGAACTGCTGCTGCA 583
Db 487 CCCTTACCTGGCCATCTCTCAAGCCCAACCAAGTCAAGGCCCAAGGAGCTGTGGCTGA 546
QY 584 AAAGGAGTCTATGTGGGCTCTGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
Db 547 AAAGGTGGTCTATGTTGGGCTCTGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 606
QY 644 CTTTGGCGAGCTCAGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCG 703
Db 607 CTTTGGCAACG-----TCAGTAGGCGAGATGACAGATATATCTGTGACCG 651
QY 704 CTTTACCCCGATAGCTGTGGATGGTGTGTTTCAATTCAGCAATATAATGGTGGGTCT 763
Db 652 CTTTACCCCGATAGCTGTGGTGGTGTGTTTCTCAGTTTCAGCACATCATGGTGGGCT 711
QY 764 CATCTGCGCGGATCGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
Db 712 TATCTGCTGCTGATTTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 771
QY 824 CTCGAAGGGCCACCAAGCGGAGCCCTCAAGACGACAGTCATCTCTCTCTCTCTCTCTCTCT 883
Db 772 CTCGAAGGGCCACCAAGCGGAGCCCTCAAGACCAAGTCATCTCTCTCTCTCTCTCTCTCT 831
QY 884 CTTTGGCTGTGCTGCCATATATATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCT 943
Db 832 CTTTGGCTGTGCTGCCATATATATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCT 891
QY 944 AGTCATCAAGCAGGATGTGACTTCGAGGAGCTGTGCAAGTGTGCAAGTGTGATCTCCATCACA 1003
Db 892 AATCATCAAGCAAGGCTGTGAGTTTGAGAACCTGTGCAAGTGTGATTTCCATCACCAG 951
QY 1004 GGCCTCGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
Db 952 GGCCTAGCTTTCTTCCATGTTGTGTGAACCCCATCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
QY 1064 GTTCAAAGCTGTGCCAGCATGCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123
Db 1012 ATTTAAACCTCTGCGCAGCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1071
QY 1124 CTTTCCAAAGGAAAGGGGGTGGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
Db 1072 CTTCTCAAAGGAAAGGGAGTGGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1131
QY 1184 TTTTCACTCCAGTAAACCTTTATGCAAGACTTATATATATATATATATATATATATATATA 1243
Db 1132 TTTTCACTCCAGTAAACAGATGTAAGACTTTTT-----TTTATACGATAAATA 1183
QY 1244 GACCTTTTATGTTACATTTTCCAGATATTAAGAGACTGACAGCTCTTGTACAGTTTT 1303
Db 1184 ACTTTTTTTAAGTTACACATTTTTTCAATATATAAAGACTGACCAATATTTGTACAGTTTT 1243

```


antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
ischemic condition; reperfusion injury; retinopathy; neonatal stress;
preclampsia; atherosclerosis; inflammatory condition; wound healing;
inflammation; erythropoiesis; hair loss; human; gene; ds.
Homo sapiens.
WO200246465-A2.
13-JUN-2002.
10-DEC-2001; 2001WO-GB005458.
08-DEC-2000; 2000GB-00030076.
08-FEB-2001; 2001GB-00003156.
25-OCT-2001; 2001GB-000025666.
(OXFO-) OXFORD BIOMEDICA UK LTD.
White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
Rayner WN;
WPI; 2002-627238/67.
Identifying a gene involved in disease for treating hypoxia-regulated
conditions, comprises comparing the transcriptome/proteome of two cell
types under different conditions and identifying a differentially
regulated gene.
Claim 37; Page 378; 538pp; English.
The present invention relates to methods for identifying genes and
proteins that are implicated in a specific disease or physiological
condition. The method comprises comparing the transcriptome/proteome of a
specialised cell type implicated in a disease or condition with that of a
second specialised cell type under two experimental conditions, and
identifying a gene that is differentially regulated in the two
specialised cell types under experimental conditions. ABV7873-ABV78116
and ABP65061-ABP65257 were identified using the methods of the invention.
The coding sequences and proteins are useful for treating hypoxia-regulated
conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
biological response to hypoxia conditions, or hypoxic-associated
pathology in a patient. The coding sequences and proteins are also useful
for monitoring the therapeutic treatment of a disease or physiological
condition, such as cancer, ischaemic conditions, reperfusion injury,
retinopathy, neonatal stress, preclampsia, atherosclerosis, inflammatory
conditions, wound healing, inflammation, erythropoiesis or hair loss
Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 50.8%; Score 952.8; DB 6; Length 1679;
Best Local Similarity 80.2%; Pred. NO. 4.1e-174;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;
QY 44 GCAGGTGCAGGTAGCAGTGCACCTCTGAGCGCTTTGGTCTCCGGTAACACCCAGCGCTG 103
Db 12 GCGGCGAGGTAGCAAGTGCAGCGCGAGGCGCTGAGTGTCTCAGTAGCCCGCATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGCATCAGTGTGAGTATATACACTTCTGATACTCTC 163
Db 72 GAGAACCGCGGTACCATGGA-----GGGGATCAGTATATACACTTACAGATACTACAC 126
QY 164 TGAAGAGTGGGTCTGGAGACTATGACTCCACACAGGAAACCTGCTCCGGGATGAAA 223
Db 127 CGAGGAAATGGGCTCAGGGAGCTATACCTCCATGAAGGAACCTGTTTCCGTGAGAAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGCCCCACCACTCTCATCATCTTCTTGACTGGCAT 283
Db 187 TGTATTTTCAATAAATCTTCTGCCCCACCACTCTCATCATCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTGGTGTGATCTGCTGATGTTGTGATGTTGTACAGAGAACTAGGAGCATGAC 343

QY 704 CTTTACCCCGATAGCCTGTGTGGATGGTGGTGTTCCTCAATTCAGCATATATATGGTGGTCT 763
Db 652 CTTTACCCCGATAGCCTGTGTGGTGGTGTTCCTCAATTCAGCATATATATGGTGGTGGTCT 711
QY 764 CATCTGCGCGCATGCTGCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTCA 823
Db 712 TATCTGCTGCTGTTGCTATCTCTGCTCTGCTATTTGATATCATCTCCAGCTGTCA 771
QY 824 CTCGAAGGCGCACAGAGCGAAGCGCCCTCAAGACGAGTGTATCTCTATCTTCTAGCTTT 883
Db 772 CTCGAAGGCGCACAGAGCGAAGCGCCCTCAAGACGAGTGTATCTCTATCTCTATCTTCT 831
QY 884 CTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943
Db 832 CTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
QY 944 AGTCATCAAGCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
Db 892 AATCATCAAGCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 951
QY 1004 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063
Db 952 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
QY 1064 GTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGAT 1123
Db 1012 ATTTAAACCTCTGCCAGCAGCACTCACTCTGTCGAGCAGAGGCTCCAGCCTCAAGAT 1071
QY 1124 CTTTCCAAAGCAAGCGGGGTGGACACTCTTCCGCTCCAGGAGTCAGATCTCTCGAG 1183
Db 1072 CCTCTCCAAAGCAAGCGGGGTGGACACTCTCTGTTTCCACTGAGTGTGAGTCTTCAAG 1131
QY 1184 TTTTCACTCCAGCTAACTTATGCAAGACTTATATATATATATATATATATATATATA 1243
Db 1132 TTTTCACTCCAGCTAACTTATGCAAGACTTATATATATATATATATATATATATATA 1183
QY 1244 GAACCTTTTATGTTTACATTTTCCAGATTAAGAGACTGACCAAGTCTTGTACAGTTT 1303
Db 1184 ACTTTTATGTTTACATTTTTCAGATATAAAGACTGACCAATATTTACAGTTT 1243
QY 1304 TTTTATTTTATGTTTACATTTTCCAGATTAAGAGACTGACCAAGTCTTGTGAGTTT 1363
Db 1244 TATTCCTGTTGGAATTTTGT-----CTGCTGTTTCTTTAGTTTGTGAGTTTATTT 1297
QY 1364 TAAATTTATATATATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1423
Db 1298 GACTTATTTATA---TAAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 1353
QY 1424 GTGGCAAGTCTTATGAGTGTGTTTATCTGTTGTGTTGTGTTGTGTTGTGTTGTGTT 1483
Db 1354 GTGGCAAGTCTTATGTTGCTGTATGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1405
QY 1484 AACTGAACATTCAGAAATGTGGTAAATTTGAATAAGCTAGCCGTGATCTCTAGCTGTT 1543
Db 1406 AACTGAACATTCAGAAATGTGGTAAATTTGAATAAGCTAGCCGTGATCTAGCTGTT 1465
QY 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473
RESULT 9
ABV78039 standard; DNA; 1679 BP.
QY ID ABV78039 standard; DNA; 1679 BP.
XX AC ABV78039;
XX AC ABV78039;
DT 12-NOV-2002 (first entry)
XX Hypoxia-regulated protein coding sequence #59.
DE Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
KW

CDS	89..1147
FT	/tag= a
FT	/product= "CXCR4"
XX	
XX	WO2003014153-A2.
XX	
XX	20-FEB-2003.
PD	
XX	12-AUG-2002; 2002WO-CA001248.
XX	
XX	10-AUG-2001; 2001US-0311088P.
PR	(TOPI-) TOPIGEN PHARM INC.
PA	
XX	Renzi P, Zemzoui K;
PI	
XX	WPI; 2003-256541/25.
DR	P-PSDB; ABP97733.
XX	
XX	Modulating viral infection of a cell, for treating or preventing
PT	respiratory virus infections, bronchitis, pneumonia or asthma, by
PT	modulating a binding interaction between a cell chemokine-receptor and a
PT	surface protein of the virus.
XX	
PS	Disclosure; Page 117-119; 120pp; English.
XX	
XX	The present sequence encodes human chemokine receptor CXCR4. The
CC	specification describes a method for modulating viral infection of a
CC	cell. the method comprises modulating a binding interaction between a
CC	cell chemokine-receptor and a surface protein of the virus. The proviso
CC	is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC	not HIV. The method is useful for treating or preventing respiratory
CC	virus infection in vertebrates, more particularly respiratory syncytial
CC	virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC	bronchitis, pneumonia or asthma
XX	
SQ	Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
	Query Match 50.8%; Score 952.8; DB 7; Length 1679;
	Best Local Similarity 80.2%; Pred. No. 4.1e-174;
	Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6
Qy	44 GCAGGTGCAGTTCGATGACCTCTGAGGGTTTGCTCGGTTAACACACGCGTG 103
Db	12 GCGGAGCAGGTAGCAAAGTAGCACGCCGAGGCCCTGAGTGTCCAGTAGCACCGCATCTG 71
Qy	104 TAGAGCGAGTGTTCGATGACCGATCAGTGTGAGTATATACATTCTGATAACTATCTC 163
Db	72 GAGAACCAGCGGTTACATGGA----GGGATCAGTATATACATTGATTAATCTACAC 126
Qy	164 TGAAGAAAGTGGGGTCTGGAGACTATGACTCCAACAGAAACCTGTCTCCGGATGAAAA 223
Db	127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGTGAAGAAA 186
Qy	224 CGTCCATTTTCATAGGACTTCCTGCCACATCTACTTCATCATCTTCTTGACTGGCAT 283
Db	187 TGCTAATTTCAATAAAA'TCTTCCTGCCACCATCTACTCCATCATCTTTTAACCTGGCAT 246
Qy	284 AGTCGGCAATGGATGGTGATCCTCGTTCATGGGTTACAGAAGAACTAAGAGCATGAC 343
Db	247 TGTGGCAATGATTTGGTCATCTCTGGTTCATGGGTTACAGAAGAACTGAGAAGCATGAC 306
Qy	344 GGACAAGTACGGCTGCACCTGTGAGTGGCTGACCTCTCTTTGTTCATCACACTCCCCTT 403
Db	307 GGACAAGTACAGGCTGCACCTGTGAGTGGCGACCTCTCTTTGTTCATCACGCTTCCCCTT 366
Qy	404 CTGGCGAGTTGATGCGCATCGGCTGACTGGTACTTTGGGAAATTTTTGTGAAGCGTGTCCA 463
Db	367 CTGGCGAGTTGATGCGCGTGGCAAACTGGTACTTTTGGGAACTTCTTATGCAAGCGTCCA 426
Qy	464 TATCATCTACATGTCAACCTCTACAGCAGCGTTCTCATCTCGGCCCTTATCATGCGCTGGA 523
Db	427 TGTCACTACAGTCAACCTCTACAGCAGTGTCTCTCATCTCGGCCCTTATCATGCTGTGA 486

RESULT 13

AB242642
ID AB242642 standard; DNA; 1679 BP.

XX AC AB242642;

XX DT 04-MAR-2003 (first entry)

XX XX Human CXK chemokine receptor 4 nucleotide SEQ ID NO:75.

DE G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
XX G protein-coupled receptor modulator; antibody; immune-related disease;
XX G growth-related disease; cell regeneration-related disease; AIDS; cancer;
XX G immunological-related cell proliferative disease; autoimmune disease;
XX G Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
XX G osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
XX G graft-versus host disease; Parkinson's disease; multiple sclerosis; pain;
XX G psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
XX G mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
XX G hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
XX G ulcer; gene; ds.

OS Homo sapiens.

XX WO200261087-A2.

XX 08-AUG-2002.

XX 19-DEC-2001; 2001WO-US050107.

XX 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burmer GC, Roush CL, Brown JP;

XX WPI; 2003-046718/04.

XX P-PSDB; ABP81796.

XX New isolated antigenic peptides e.g., for G protein-coupled receptors
XX (GPCR), useful for diagnosing and designing drugs for treating conditions
XX in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
XX autoimmune diseases.

XX Disclosure; Fig 1; 533pp; English.

XX The present invention describes antigenic peptides (I) comprising: (a)
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
XX acids. Also described: (1) an assay for the detection of a particular G
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
XX and (2) an isolated antibody having high specificity and high affinity or
XX avidity for a particular GPCR. (I) can be used as GPCR modulators and in
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an
XX antibody against a particular GPCR, and in the production of specific
XX antibodies. The peptides and antibodies are also useful for detecting the
XX presence or absence of corresponding GPCRs. The antigenic peptides for
XX GPCRs and antibodies are useful for diagnosing and designing drugs for
XX treating immune-related diseases, growth-related diseases, cell
XX regeneration-related disease, immunological-related cell proliferative
XX diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
XX atherosclerosis, bacterial, fungal, protozoan or viral infections,
XX osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
XX inflammation, allergies, Crohn's disease, diabetes, graft versus host
XX disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
XX anxiety, depression, schizophrenia, dementia, mental retardation, memory
XX loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, or
XX hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
XX any other disorder in which GPCRs are involved. The antibodies may be
XX used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
XX GPCR proteins given in ABP81675 to ABP82018, which are used in the
XX exemplification of the present invention

XX

SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 50.8%; Score 952.8; DB 7; Length 1679;
Best Local Similarity 80.2%; Pred No. 4.1e-174;
Matches 1210; Conservative 0; Mismatches 252; Indels 45; Gaps 6;
QY 44 GCAGGTGCAGGTAGCAGTGAACCTCTGAGGCGTTGGTCTCGGTAAACACACCGGTG 103
DB 12 GCGGCAGCAGGTAGCAAGTGAACGCGAGGGCTGAGTCTCCAGTAGGACCGCATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGCATCAGTGTGAGTATATACACTTCTGATAACTACTC 163
DB 72 GAGAACAGCGGTACATGGA-----GGGATCAGTATATACACTTCAGATAACTACAC 126
QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTTCAACAGAACCTCTGCTCCGGGATGAAAA 223
DB 127 CGAGGAATAGGGCTCAGGAGCACTATGACTCCATGAAGAACCTCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGATCTTCTGCGCCACCACTACTTTCATCATCTTCTTGACTGGCAT 283
DB 187 TGTAAATTTCAATAAATCTTCTGCGCCACCACTACTTTCATCATCTTCTTGACTGGCAT 246
QY 284 AGTCGGCAATGGATTGGTGTATCCTGTCTATGGGTTACCAAGAAAGACTAAGAGCATGAC 343
DB 247 TGTGGCAATGGATTGGTGTATCCTGTCTATGGGTTACCAAGAAAGACTAAGAGCATGAC 306
QY 344 GGAACAGTACCGCTGCACCTGTCTAGTGGCTGACCTCTTGTCTCATCACCTCCCTT 403
DB 307 GGACAAAGTACAGGCTGCACCTGTCTAGTGGCGGACCTCTCTTTGTCTATCAGCTTCCCTT 366
QY 404 CTGGCGAGTTGATGCGCATGGCTGACTGCTTCTTGGGAAATTTTGTGAAGCTGTCCA 463
DB 367 CTGGCGAGTTGATGCGCTGGCAACTGTTGTTGGAACTTCTTATGCAAGCAGTCCA 426
QY 464 TATCATCTACACTGTCACTCTACAGAGCGTTCTCATCTCTGGCTTCTATCAGCGTGA 523
DB 427 TGTCACTACACAGTCAACCTCTACAGAGCTGTCCTCTCTGCTTCTATCAGTCTGGA 486
QY 524 CCGGTACTCTGCCATTGTCCACGCCCAACCAAGCTCAAGGCCAAGGAACTGCTGGCTGA 583
DB 487 CCGGTACTCTGCCATTGTCCACGCCCAACCAAGCTCAAGGCCAAGGAGCTGTTGCTGA 546
QY 584 AAAGCGAGTCTATGTGGGCGTCTGGATCCAGCGCTCTCTCTGACTATATCTGACTTCAT 643
DB 547 AAAGGTGGTCTATGTGGCGTCTGGATCCCTGCGCTCTCTGCTGACTATTCCTGACTTCAT 606
QY 644 CTTTGGCGAGTCTGAGCGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCG 703
DB 607 CTTTGGCGAGTCTGAGCGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCG 651
QY 704 CTTTGGCGAGTCTGAGCGGGGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 763
DB 652 CTTTGGCGAGTCTGAGCGGGGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 711
QY 764 CATCTGCGCGGATCTGATCT 823
DB 712 TATCTCTGCTGATCTGATCT 771
QY 824 CTCGAAGGCCACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 883
DB 772 CTCGAAGGCCACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 831
QY 884 CTTTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
DB 832 CTTTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
QY 944 AGTCATCAAGCAAGGATGTGACTTCCGAGAGCATTGTGCAAGTGGATCTCCATCACAGA 1003
DB 892 AATCATCAAGCAAGGATGTGAGTTTGAGACACTGTGCAAGTGGATTTCCATCACCGA 951
QY 1004 GGCGCTCGCTCTCTTCCACTGTTGCCGTGAACCCCATCTCTATGCGCTTCTCTCGGGGCAA 1063
DB 952 GGCGCTAGCTTTCTTCCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1011

QY 1064 GTTCAAGGCTCTCCAGCATGCACTCACTCCATGAGCAGAGGCTCCAGGCTCAAGAT 1123
DB |||
QY 1012 ATTAAACCTCTGCCAGCAGCAGCTCACTCTGTGAGCAGAGGCTCCAGGCTCAAGAT 1071
DB |||
QY 1124 CTTTCCAAAGGAAGCGGGTGACACTCTTCGGTCTCCACGAGTCAAGATCTCCAG 1183
DB |||
QY 1072 CTTCTCAAAGGAAGCGAGGTGACATTCATCTGTTTCCACTGAGTCTGAGTCTCAAG 1131
DB |||
QY 1184 TTTTCACTCCAGCTAACCTTATGCAAGACTTATATAATATATATATATATATA 1243
DB |||
QY 1132 TTTTCACTCCAGCTAACAGATGTAAGACTTTT-----TTTACGATTAATA 1183
DB |||
QY 1244 GAACCTTTTATGTATACACATTTTCCAGATATAAGAGACTGACCACTTTGTACAGTTT 1303
DB |||
QY 1184 ACTTTTAAAGTTACACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTT 1243
DB |||
QY 1304 TTTTCTTTTAAATGACTGTGGAGTTTATGTTCTCTAGTTTGTGAGGTTTGACT 1363
DB |||
QY 1244 TATTGCTGTGGATTTTGT-----CTTGTGTTCTTTAGTTTGTGAAAGTTAAT 1297
DB |||
QY 1364 TAATTTATATAAATATGTTTTTGTGTTTCTATGTGAATGAGCGTCTAGGCGACCT 1423
DB |||
QY 1298 GACTTATTTATA-----TAATTTTGTGTTTCTATGATGTTGTTAGGCGACCT 1353
DB |||
QY 1424 GTGCCAAGTTCTTAGTGTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGGA 1483
DB |||
QY 1354 GTGCCAAGTTCTTAGTGTGTTTATCTGTGTAGGACTGTAGAA-----AAGGG 1405
DB |||
QY 1484 AACTGACATTCAGATGTGTGTAATTAATGAATAAGCTAGCGGTGATCCTCAGCTGT 1543
DB |||
QY 1406 AACTGACATTCAGAGCGGTGTAGTGAATCACGTAAAGCTAGAAATGCCAGCTGT 1465
DB |||
QY 1544 GCTGCATA 1551
DB |||
QY 1466 TATGCATA 1473
DB |||

RESULT 14

ID ADC98645
XX ADC98645 standard; cDNA; 1679 BP.

AC ADC98645;

XX 01-JAN-2004 (first entry)

XX Human CXc chemokine receptor 4 (CXCR4)-encoding cDNA.

XX Human; CXc chemokine receptor 4; CXCR4; CXCR4 inhibitor;
XX small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
XX metastasis; motility; morphological change; drug screening; monitoring;
XX cytosolic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

PH 89..1147

XX CDS /**tag= a

FT /product= "Human CXCR4"

FT WO2003079020-A2.

XX 25-SEP-2003.

XX 20-MAR-2003; 2003WO-EP002916.

XX 20-MAR-2002; 2002US-0366370P.

XX (DAND) DANA FARBER CANCER INST INC.

XX (NOVS) NOVARTIS AG.

XX Salgia R;

XX

DR WPI; 2003-757037/71.

DR P-PSDB; ADC98646.

XX Use of a CXc chemokine receptor 4 (CXCR4) inhibitor, for modulating
PT cellular adhesion or inhibiting proliferation, movement, or morphological
PT change in a small cell lung cancer (SCLC) cell population, or for
PT treating a subject having SCLC.

XX Disclosure; SEQ ID NO 1; 78pp; English.

XX The invention relates to methods and compositions for inhibiting small
CC cell lung cancer (SCLC) proliferation and metastasis through modulation
CC of the activity or expression of CXc chemokine receptor 4 (CXCR4), and
CC optionally, that of the tyrosine kinase receptor c-Kit. Administration of
CC a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
CC proliferation, motility, and morphological changes in an SCLC cell
CC population. The invention is based on the finding that CXCR4 is
CC ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
CC The invention also discloses methods of screening for CXCR4 inhibitors;
CC determining whether a sample of lung cancer cells expresses CXCR4 and is
CC therefore amenable to CXCR4 inhibitor therapy; determining whether a
CC sample of lung cancer cells coexpress CXCR4 and c-Kit is is therefore
CC amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
CC inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
CC CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
CC by determining CXCR4 (and optionally c-Kit) activity in two or more
CC tumour samples taken from an SCLC patient during the course of treatment.
CC The methods of the invention are useful for treating SCLC, for screening
CC for agents for the treatment of SCLC, for determining whether a SCLC from
CC individual patients is suitable for treatment with agents of the
CC invention, and for determining whether such treatment is effective and
CC should be continued. The present sequence represents cDNA encoding human
CC CXCR4, which can be used in methods of the invention.

XX Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

XX Query Match 50.8%; Score 952.8; DB 9; Length 1679;

XX Best Local Similarity 80.2%; Pred. No. 4.1e-174;

XX Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

QY 44 GCAGGTGAGGTAGCAGTGAACCTCTGAGGCGTTTGTGTCGCGTAAACCACCGGCTG 103

DB ||| 12 GCGGCGAGGTAGCAAAAGTGAACGCGGAGGCGCTGAGTGCTCCAGTACCCACCGCATCTG 71

QY 104 TAGAGCGAGTGTGCGCATGGAACCGATCAGTGTGATATATACATCTCTGTAACCTACTC 163

DB ||| 72 GAGAACCGCGGTACCATGGA-----GGGATCAGTATATACATCTCAGATTAATACAC 126

QY 164 TGAAGAAGTGGGCTCTGGAGACTATGACTCCCAAGAGAACCTGCTTCCGGGATGAAAA 223

DB ||| 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCTGTTTCCGTGAAGAAAA 186

QY 224 CGTCCATTTCAATAGGATCTTCTGCGCCACCATCTACTTCTATCATCTTTTGGACTGGCAT 283

DB ||| 187 TGCTAATTTCAATAAAATCTTCTGCGCCACCATCTACTTCTTCTTAACTGGCAT 246

QY 284 AGTCGGCAATGGATTGGTGATCTCTGCTCATGGGTTTACCAGAAGAGCTTAAGGAGCATGAC 343

DB ||| 247 TGTGGCAATGGATTGGTGATCTCTGCTCATGGGTTTACCAGAAGAGCTTAAGGAGCATGAC 306

QY 344 GGACAAGTACCGGCTGCACCTGTGAGTGGGTGACCTCTCTTTTGTTCATCACTCCCTCCCT 403

DB ||| 307 GGACAAGTACAGGCTGCACCTGTGAGTGGGTGACCTCTCTTTTGTTCATCACTCCCTCCCT 366

QY 404 CTGGGCGAGTTGATGCCATGGCTGAGTGGTACTTTTGGGAAATTTTGTGTAAGGCTGTCCA 463

DB ||| 367 CTGGGCGAGTTGATGCCATGGCTGAGTGGTACTTTTGGGAAATTTTGTGTAAGGCTGTCCA 426

QY 464 TATCATCTACACTGTCAACCTTCTACAGCAGCGTTCCTCTCTCTGGCTTCTATCAGCTTGA 523

DB ||| 427 TGTCTATCTACAGTCAACCTTCTACAGCAGTGTCTCTCTCTGGCTTCTATCAGCTTGA 486

QY 524 CCGGTACCTCGCCATGTGCCAGCGCCACCAACAGTCAAGGCCCAAGGAACCTCTGCTGA 583

Db 487 CGCTACCTGGCCATGTCACGCCACCACAGTCAGAGGCCAAGAAAGCTGTTGGCTGA 546
Qy 584 AAAGGCACTATGTGGGGTCTGGATCCAGCCCTCTCTCTGACATATACCTGACTTCAT 643
Db 547 AAAGTGGTCTATGTGGGCTGGATCCCTGCCCTCTCTGCTGACTATTCGGACTTCAT 606
Qy 644 CTTTGGCAGCTGACGCCAGGGGACATCAGTCAGGGGAGTACAGGTACATCTGACGG 703
Db 607 CTTTGGCAACG-----TCAGTGAGCAGATGACATATATCTGFGACGG 651
Qy 704 CTTTACCCTGATGAGCTGTGGATGGTGTGTTTCAATTCAGCATATAATGGTGGTCT 763
Db 652 CTTTACCCTGATGACTTGTGGGTGGTGTGTTCCAGTTCAGCACATCAATGGTGGCT 711
Qy 764 CATCTGCGGGATCGTCATCTCTCTGTTACTGTCATCATCTCTGAGCTGTACA 823
Db 712 TATCTGCTGCTGATTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 771
Qy 824 CTCACAGGGCCACAGAGCGGAGCCCTCAGAGCAGTCCCTCATCTCTGCTGCTT 883
Db 772 CTCACAGGGCCACAGAGCGGAGCCCTCAGAGCAGTCCCTCATCTCTGCTGCTT 831
Qy 884 CTTTGGCTGCTGCTGCCATTTATGTGGGGATCAGCATCGCTCTCTCATCTCTGCTTGG 943
Db 832 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
Qy 944 AGTCATCAAGCAGGATGTGACTTCGAGAGCATTTGCAAGTGATGCTCCATCACAGA 1003
Db 892 AATCATCAAGCAGGATGTGAGTTGAGAACACTGTGCAAGTGATGCTCCATCACAGA 951
Qy 1004 GGCCTGCTGCTTCTTCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
Db 952 GGCCTGCTGCTTCTTCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1011
Qy 1064 GTTCAAAGCTGCTGCCAGATGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 1123
Db 1012 ATTAAAGCTGCTGCCAGCAGCTCACTCACTCACTCACTCACTCACTCACTCACTCA 1071
Qy 1124 CTTTCCAAAGGAAGGGGGTGGACATCTTCCGCTCTCCAGGAGTCAAGATCTCTCCAG 1183
Db 1072 CTTTCCAAAGGAAGGGGGTGGACATCTTCCGCTCTCCAGGAGTCAAGATCTCTCCAG 1131
Qy 1184 TTTTCACTCAGCTAACCTTTATGCAAGCTTATATATATATATATATATATATATATAT 1243
Db 1132 TTTTCACTCAGCTAACAGATGTAAAGACTTTT-----TTTATCAATAAATA 1183
Qy 1244 GAACCTTTTATGTTACACATTTTCCAGATATAGAGACTGACCACTCTGTACAGTTT 1303
Db 1184 ACTTTTATTTAGTTACACATTTTCCAGATATAGAGACTGACCACTCTGTACAGTTT 1243
Qy 1304 TTTTATTTTATTTAGTTGCTGGAGTTATCTTCTCTAGTTTGTGGAGTTTGTACT 1363
Db 1244 TATTGCTTGTGGATTTTGT-----CTTGTCTTTCTTTAGTTTGTGGAGTTTAAAT 1297
Qy 1364 TAAATTTATATATATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1423
Db 1298 GACTTATTTATA-----TAAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1353
Qy 1424 GTGGCCAAAGTCTTAGTAGCTGTTTATCTGTGTAGGAGCTGTAGAACTGTAGAGAGA 1483
Db 1354 GTGGCCAAAGTCTTAGTTGCTGTATGCTGTGTAGGAGCTGTAGAA-----AAGGG 1405
Qy 1484 AACTGAACATTCAGAAATGTGGTAAATTAAGTACGCTGAGCTGATCTCAGCTGTT 1543
Db 1406 AACTGAACATTCAGAGCGGTGTAGTGAATCAGCTAAAGCTAGAAATGATCCCGAGCTGT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473

RESULT 15

ABZ35348
ID ABZ35348 standard; cDNA; 1711 BP.
XX AC ABZ35348;
XX DT 05-FEB-2003 (first entry)
XX Human gene expression profile polynucleotide SEQ ID NO 459.
XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
XX bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
XX tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
XX gene expression; gene; ss.
XX Homo sapiens.
XX WO200274979-A2.
XX 26-SEP-2002.
XX 20-MAR-2002; 2002WO-US008456.
XX 20-MAR-2001; 2001US-0276947P.
XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX Wan J, Wang Y;
XX WPI; 2002-740862/80.
XX New gene expression profile generated from primary, endothelial,
XX epithelial, and muscle cell types, useful for identifying disease
XX pathologies involving alterations of gene expression, e.g. cancer.
XX Disclosure; Page 609; 850pp; English.
XX The invention relates to a gene expression profile comprising one or more
XX genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type
XX is a coronary artery endothelium, umbilical artery or vein endothelium,
XX aortic endothelium, dermal microvascular endothelium, pulmonary artery
XX endothelium, myometrium microvascular endothelium, keratinocyte
XX epithelium, bronchial epithelium, mammary epithelium, prostate
XX epithelium, renal cortical epithelium, renal proximal tubule epithelium,
XX small airway epithelium, renal epithelium, umbilical artery smooth
XX muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
XX dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
XX aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
XX bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
XX osteoblasts or prostate stromal cell. The gene expression profile is used
XX for determining the level of RNA expression for a sample, determining the
XX phenotype of a cell and distinguishing cell types. The gene or a protein
XX expression profile is useful in identifying disease pathologies involving
XX alterations of gene expression. The assessment of expression profiles may
XX provide meaningful information with respect to tumour type and stage,
XX treatment methods, and prognosis. The gene or protein expression profile
XX may also be used for creating microarrays. The microarray is useful for
XX genetic and physical mapping of genomes, DNA sequencing, genetic or
XX medical diagnosis, genotyping of organisms, confirming cell or tissue
XX identifications and in identifying promising antibiotics, antiviral or
XX antifungal agents
XX Sequence 1711 BP; 423 A; 407 C; 373 G; 507 T; 0 U; 1 Other;

Query Match 50.8%; Score 952.8; DB 6; Length 1711;
Best Local Similarity 80.2%; Pred. No. 4.2e-174;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

Qy 44 GCAGTGCAGGTAGCAGTGACCCCTCTCAGGCGTTTGTGCTCCGGTAACCAACCGCTG 103
Db 47 GCGGCAGCAGGTAGCAAAAGTGACCGCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCTG 105
Qy 104 TAGAGCGAGTGTGCCATGGACCGATCAGTGTGAGTATATACACTTCTGATACTACTC 163

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 150.281 Seconds
(without alignments)
6931.281 Million cell updates/sec

Title: US-09-367-052-1

Perfect score: 1977

Sequence: 1 ccatcctaatacagactcact.....aaaaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCFUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	974.4	51.9	1737	1	US-08-202-056-4
2	974.4	51.9	1737	1	US-08-076-092A-3
3	974.4	51.9	1737	1	US-08-701-285-3
4	974.4	51.9	1737	2	US-08-284-586-3
5	974.4	51.9	1737	2	US-08-805-478-3
6	974.4	51.9	1737	2	US-08-802-627A-3
7	974.4	51.9	1737	2	US-08-801-238-3
8	974.4	51.9	1737	2	US-08-801-228-3
9	974.4	51.9	1737	3	US-09-104-296-3
10	974.4	51.9	1737	5	PCT-US94-06380-2
11	952.8	50.8	1679	4	US-09-517-605-14
12	951.8	50.7	1664	4	US-09-582-224A-5
13	951.8	50.7	1664	4	US-09-023-655-1213
14	859	45.8	1225	4	US-09-016-434-1235
15	834.6	44.5	1317	1	US-08-133-848-45
16	834.6	44.5	1317	3	US-09-239-843A-45
17	834.6	44.5	1317	4	US-09-088-337B-45
18	834.6	44.5	1317	5	PCT-US93-11153-45
19	198.4	10.6	1119	4	US-09-170-496D-65
20	198.4	10.6	1679	1	US-08-202-056-6
21	198.4	10.6	1679	1	US-08-076-093A-5
22	198.4	10.6	1679	1	US-08-701-285-5
23	198.4	10.6	1679	2	US-08-284-586-5
24	198.4	10.6	1679	2	US-08-805-478-5
25	198.4	10.6	1679	2	US-08-802-627A-5
26	198.4	10.6	1679	2	US-08-801-238-5
27	198.4	10.6	1679	2	US-08-801-228-5

28 198.4 10.6 1679 3 US-09-104-296-5
29 198.4 10.6 1679 5 PCT-US94-06380-3
30 198.4 10.6 2818 3 US-08-982-493-7
31 198.4 10.6 2818 3 US-08-628-655-1
32 196.8 10.5 1119 4 US-09-170-496D-199
33 195.8 10.4 2085 3 US-09-299-843A-65
34 195.8 10.4 2085 4 US-09-088-337B-65
35 193.8 10.3 1107 4 US-09-170-496D-19
36 193.8 10.3 1293 4 US-09-016-434-1052
37 193.8 10.3 1670 3 US-08-709-839-1
38 193.8 10.3 1670 3 US-08-829-839-1
39 193.8 10.3 1670 4 US-09-023-655-980
40 193.8 10.3 1670 4 US-09-624-594-1
41 192.2 10.2 1107 4 US-09-170-496D-173
42 186.8 10.0 2751 1 US-08-153-848-23
43 186.8 10.0 2751 3 US-09-299-843A-23
44 186.8 10.0 2751 4 US-09-088-337B-23
45 186.8 10.0 2751 5 PCT-US93-11153-23

ALIGNMENTS

RESULT 1
US-08-202-056-4
; Sequence 4, Application US/08202056
; Patent No. 5440021
; GENERAL INFORMATION:
; APPLICANT: Chuncharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,056
FILING DATE: 25-FEB-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-202-056-4

Query Match 51.9%; Score 974.4; DB 1; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

QY 50 GCAGGTAGCAGTGAACCTCTGAGGCGTTTGTGTCTCCGGTAACCAACCGGCTGTAGAGC 109
Db 20 GCGCGCGCAAGTGAAGCGCGAGGCGCTGAGTGTCTCCAGTAGGACCAACCGCATCTGAGAGC 79
QY 110 GAGTGTGCGATGGAACCGATCAGTGTGAGTATATACACTTCTGATTAATCTACTCTGAGA 169
Db 80 CAGCGGTTACCATGGA-----GGGATCAGTATATACACTTCTGATTAATCTACTCTGAGA 134
QY 170 AGTGGGCTCTGAGAGTATGACTCCAAAGAAACCTCTCTCCGGGATGAAGAGTCCCA 229
Db 135 AATGGGCTCAGGGACTATGACTCCATGAGAAACCTCTCTCCGGTGAAGAAATGCTAA 194
QY 230 TTTCAATAGGATCTCTGCGCCCAACCATCTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 289
Db 195 TTTCAATAAATCT 254
QY 290 CAATGATGTTGATCT 349
Db 255 CAATGATGTTGATCT 314
QY 350 GTACCGGCTGCACT 409
Db 315 GTACCGGCTGCACT 374
QY 410 AGTTGATGCGATGCT 469
Db 375 AGTTGATGCGATGCT 434
QY 470 CTACACTGTCAACCT 529
Db 435 CTACACTGTCAACCT 494
QY 530 CTTGCGGCTGCACT 589
Db 495 CTTGCGGCTGCACT 554
QY 590 AGTCTATGTTGGGCTCTGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649
Db 555 GGTCTATGTTGGGCTCTGATCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 614
QY 650 CGAGCTGACGAGGGGCAATCAGTCAGGGGAGTACAGGTACATCTGTGACCGCTTTA 709
Db 615 CAACG-----TCAGTGAGGAGTACAGATATATCTGTGACCGCTTTA 659
QY 710 CCCCAGATGCTGTGATGTTGTTCAATTCAGGATATAATGTTGGGCTCTATCTCT 769
Db 660 CCCCAGATGCTGTGATGTTGTTCCAGTTCAGCACATCATGTTGGGCTTTATCTCT 719
QY 770 GCGCGGATGCTATCT 829
Db 720 GCGCGGATGCTATCT 779
QY 830 GCGCGGATGCTATCT 889
Db 780 GCGCGGATGCTATCT 839
QY 890 CTGCTGGCTGCCATATATGTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCT 949
Db 840 CTGTTGGCTGCCATATATGTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCTCT 899
QY 950 CAAGCAAGGATGCTATCT 1009
Db 900 CAAGCAAGGATGCTATCT 959
QY 1010 CGCT 1069
Db 960 AGCTTCT 1019
QY 1070 AGCTTCT 1129
Db 1020 AACCT 1079
QY 1130 CAAAGGAAAGCGGGTGGACACT 1189

Db 1080 CAAAGGAAAGCGAGGTGGACATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1139
QY 1190 CTCAGCTAACCTTATGCAAGACTTATATATATATATATATATATATATATATATATATATAT 1249
Db 1140 CTCAGCTAA-----CACAGATCTAAAGACTTTTTTTTATACGATAAATAACTTTT 1191
QY 1250 TTTTATGTTACATTTTCCAGATATAAGAGACTGACCACTCTCTCTCTCTCTCTCTCTCTCTCT 1309
Db 1192 TTTTATGTTACATTTTCCAGATATAAAGACTGACCACTTATGTTTATGCTTTTATGCT 1251
QY 1310 TTTTATGTTACATTTTCCAGATATAAAGACTGACCACTTATGTTTATGCTTTTATGCTTT 1369
Db 1252 TTTTATGTTACATTTTCCAGATATAAAGACTGACCACTTATGTTTATGCTTTTATGCTTT 1305
QY 1370 ATATAATATGTTTCT 1429
Db 1306 TTTTATA-----TAAATTTTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1361
QY 1430 AAGTCTCTAGTACTGTTTATCTGTGTAGGACTGTAGAACTGTAGAGAAAGAACTGA 1489
Db 1362 AAGTCTCTAGTACTGTTTATCTGTGTAGGACTGTAGAACTGTAGAACTGTAGAACTGA 1413
QY 1490 ACATTCAGAAATGTTGTTAAATTAAGCTAGCCGCTGATCTCTCTCTCTCTCTCTCTCTCTCT 1549
Db 1414 ACATTCAGAGCGTGTAGTGAATCAGCTAAAGCTAGAAATGATCCCGCAGCTGTTATGCA 1473
QY 1550 TAATCT 1609
Db 1474 TAGAATCT 1488
QY 1610 TGTGTTGTTATGCTGTGTAGTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1669
Db 1489 -----TTCGGTGGAACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1525
QY 1670 TTTTCTGTAAAGATGCGACTTAAACCAAGCTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1729
Db 1526 TTTGCTGTAGAGATGCGACTTAAACCAAGCTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1581
QY 1730 TTTTCTGTAAAGATGCGACTTAAACCAAGCTGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1789
Db 1582 -----TTCGGTGGAACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1623
QY 1790 GTACAGCTCTGTATACATTTGTTTAAAGCTCAATGATTAACCTTAAAGCTCAATGATTAAC 1849
Db 1624 GTACAGCTCTGTATACATTTGTTTAAAGCTCAATGATTAACCTTAAAGCTCAATGATTAAC 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 2

US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM: 1.44 Mb floppy disk
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/076.093A
 FILING DATE: 11-Jun-1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/810782
 FILING DATE: 19-DEC-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/677211
 FILING DATE: 29-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Love, Richard B
 REGISTRATION NUMBER: 34,659
 REFERENCE/DOCKET NUMBER: 706P2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-5530
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1737 nucleotides
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 US-08-076-093A-3

Query Match 51.9%; Score 974.4; DB 1; Length 1737;
 Best Local Similarity 76.1%; Pred. No. 2.3e-195;
 Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

QY	50	CGAGGTACAGTGCCTCTGAGCGGTTTGGTGTCCGGTAACACACACGCGCTGTAGAGC	109
Db	20	CGCGGGCGAAGTCAGCGCCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCTGGAGAAC	79
QY	110	GAGTGTGCCATGGAACCGATCAGTGTGAGTATATACATCTCTGATTAATCTCTGAAGA	169
Db	80	CAGCGGTTACCATGGA-----GGGGATCAGTATATACATCTCAGATAACTACACCGAGGA	134
QY	170	AGTGGGGTCTGGAGACTATGACTCCAAACAGGAACCTGCTTCGGGATGAAGAAACGTCCA	229
Db	135	AATGGGCTCAGGGAGACTATGACTCAGTGAAGGAACCTGTTCCGTGAAGAAATGCTAA	194
QY	230	TTTCAATAGGATCTTCCTGCCACCATCTACTTATCATCTCTCTTGACTGGCATAGTCGG	289
Db	195	TTTCAATAAATCTTCCTGCCACCATCTACTCCATCATCTTCTTAACCTGGCATTTGGGG	254
QY	290	CAATGGATTTGATCTCTGTCATGGTTACCAAGAGAGCTAGGAGCATGACGGACAA	349
Db	255	CAATGGATTTGATCTCTGTCATGGTTACCAAGAGAACTGAGAGCATGACGGACAA	314
QY	350	GTACCGGTGCACCTGTGAGTGGCTGACCTCTCTTTGTATCATCACTACCTCCCTTCTGGGC	409
Db	315	GTACAGGCTGCACCTGTGAGTGGCGGACCTCTCTTTGTATCATCACTCTCCCTTCTGGGC	374
QY	410	AGTTGATGCCATGGCTGACTGGTACTTTGGAAATTTTGTGAAGCTGTCCATATCAT	469
Db	375	AGTTGATGCCATGGCTGACTGGTACTTTGGAAATTTTGTGAAGCTGTCCATATCAT	434
QY	470	CTACACTGTCAACTCTACAGCAGCGTTCTCATCTCGGCCCTTCATCAGCCTGGACCGGTA	529
Db	435	CTACACAGTCAACTCTACAGCAGTGTCTCATCTCGGCCCTTCATCAGTCTGGACCGCTA	494
QY	530	CCTGCCATTTGTCCAGCCCAACAAGTCAAGGCCAAGGAACTGTGCTGTGAAGAGGC	589
Db	495	CCTGCCATTCGTCCAGCCCAACAAGTCAAGGCCAAGGAACTGTGCTGTGAAGAGGT	554
QY	590	AGTCTATCTGGCGCTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTTCATCTTTTC	649
Db	555	GGTCTATCTGGCGCTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTTCATCTTTTC	614
QY	650	CGACGTACGCCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCCTTA	709

Db	615	CAACG-----TCAGTGAAGCAGATGACAGATATATCTGTGACCGCTTCTA	659
QY	710	CCCCGATAGCCTGTGGATGGTGGTGTTCATTCAGCATATAAATGGTGGGTCTCATCT	769
Db	660	CCCCAATGACTTGTGGTGGTGGTGTTCAGTTCAGCACATCATGGTGGCTTATCTCT	719
QY	770	GCCGGCATCTCATCTCTCTCTGTACTGTGATCATCTCTTAAGCTGTCACTCCAA	829
Db	720	GGCTGGTATTGTTCATCTCTCTCTGTCTTATGCAATATCATCTCCAAGCTGTCACTCCAA	779
QY	830	GGGCCACCAAGACGCAAGGCCCTCAAGACGACAGTCACTCTCATCTCTAGCTTTCTTTC	889
Db	780	GGGCCACCAAGACGCAAGGCCCTCAAGACACAGTCACTCTCATCTCTAGCTTTCTTTC	839
QY	890	CTGCTGGCTGCCATATATATGTGGGATCAGATCGACTCCTTCATCTCTTTCGGAGTCA	949
Db	840	CTGTTGGCTGCCCTTACTACATTTGGGATCAGATCGACTCCTTCATCTCTCTGGAATCAT	899
QY	950	CAAGCAAGGATGTGACTTCGAGAGCAATTTGTGCACAAGTGGATCTCCATCACAGAGCCCT	1009
Db	900	CAAGCAAGGATGTGAGTTTGAGAACACATGTGCACAAGTGGATTTCCATCACAGAGCCCT	959
QY	1010	CGCTTCTTCCACTGTGCTGAAACCCCATCTCTATGCCCTCTCTCGGGGCCAAGTTCAA	1069
Db	960	AGCTTCTTCCACTGTGCTGAAACCCCATCTCTATGCTTTCTCTGGAGCCAAATTTAA	1019
QY	1070	AAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCCTTTC	1129
Db	1020	AACCTCTGCCAGCAGCACTCACTCTGTGAGCAGAGGGTCCAGCTCAAGATCCTCTC	1079
QY	1130	CAAGGAAGACGGGTGGACACTCTTCGTCTCCACGGAGTCAGAAATCTCTCCAGTTTTCA	1189
Db	1080	CAAGGAAGACGGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAGTTTTCA	1139
QY	1190	CTCCGCTTAACCTTATGCAAGACTTATATAATATATATATATATATATATATATATAT	1249
Db	1140	CTCCAGCTAA-----CACAGATGTAAGACTTTTTTTTATGATTAATAACTTTT	1191
QY	1250	TTTTATGTTACATTTTCCAGATATAAGAGACTACCAAGTCTGTACAGTTTTTTTTT	1309
Db	1192	TTTTAAGTTTACATTTTTCAGATATAAAGACTACCAATATTGTACAGTTTTTATTTC	1251
QY	1310	TTTTTAATGACTGTGGAGTTTATGTCCTCTAGTTTTTTTGTGAGTTTGTACTTAATTT	1369
Db	1252	TTGTTGGATTTTGT-----CTTGTGTTTCTTTAGTTTTTTGTGAAGTTTTTAATGACTTA	1305
QY	1370	ATATAAATATTGTTTTTTTGTGTTTTCATGTAAGCTCTAGGCGAGCACTGTGGCC	1429
Db	1306	TTTATA-----TAAATTTTTTTTGTGTTTTCATATTGATGTGTCTAGCGAGCACTGTGGCC	1361
QY	1430	AAGTCTCTAGTGTGTTTATCTGTGTGAGACTGTAGAACTGTAGAGGAAGAACTGA	1489
Db	1362	AAGTCTCTAGTGTGTTTATCTGTGTGAGACTGTAGAACTGTAGAGGAAGAACTGA	1413
QY	1490	ACATTCACAAATGTGGTAAATTAATAAAGCTAGCGTGTATCTCAGCTGTGCTGA	1549
Db	1414	ACATTCACAGCGGTGTAGTGAATCAGTAAAGCTAGAAATGATCCCCAGCTGTATGCA	1473
QY	1550	TAATCTCTTTCATTCGAGGAGCACCCCAACCCCAACCCCAACCCCAACCCCAATCTTAAAT	1609
Db	1474	TAGATAATCTCTCCA-----TTCCCGTGAAACGTTTTTCTCTGTTCTTTAAGACGTGAT	1488
QY	1610	TGTTTGGTTATCTGTGTGATGGTTTTTGTGTTTTTTTGTGTTGTTGTTTTTTT	1669
Db	1489	-----TTCCCGTGAAACGTTTTTCTCTGTTCTTTAAGACGTGAT	1525
QY	1670	TTTTCTGTAAGAGTGGCACTTAAACCAAGCCCTGAAATGGTGGTAGAAATGCTGGGT	1729
Db	1526	TTTGTGTAGAGAGTGGCACTTAAACCAAGCCCAAGAGTGGT-ATAGAAATGCTGG---	1581
QY	1730	TTTTTTTGTGTTGTTTTTTTTCAGTTTTCAGAGTAGATTGACTTCAGTCCCTCAAAAT	1789

Db 1582 -----TTTTTCAGTTTTCAGGAGTGGGTTGATTTCAGCACTAC-AGT 1623
Qy 1790 GTACAGTCTGTATACATGTTTATARAAGTCATGATAAATCTTAAAAAAGAAAAA 1849
Db 1624 GTACAGTCTGTATTAAGTTGTTTAAATAAAGTACATGTTTAAACTTTAAAAAAGAAAAA 1683
Qy 1850 AAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 3

US-08-701-265-3
; Sequence 3, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,265
; FILING DATE: 22-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-Jun-1993
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-701-265-3

Query Match 51.9%; Score 974.4; DB 1; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;
Qy 50 GCAGGTAGAGTACCCCTCTGAGCGCTTGGTGCTCCGGTAAACCACCGGCTGTAGAC 109
Db 20 GCAGCGCGCAAGTGAAGCGCGCGGCTGAGTCTCCAGTAGCCACCGCATCTGGAGAC 79
Qy 110 GAGTGTGGCATGGAACCGCATCAGTGTGAGTATATACATCTCTGATAAATCTCTGAAGA 169
Db 80 CAGCGGTACCATGGA-----GGGATCAGTATATACATCTCAGATAACTACACCGAGA 134

Qy 170 AGTGGGTCTGGAGACTATGACTCCAAAGAACCCCTGCTCCGGGATGAAACGTCOA 229
Db 135 AATGGGCTCAGGGAGATATGACTCCATGAGGAAACCCCTGTTCCGTGAGAGAAATGCTAA 194
Qy 230 TTTCAATAGAGATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTTCAGTGGCATAGTCGG 289
Db 195 TTTCAATAAATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTAACTGGCATTTGGG 254
Qy 290 CAATGGATTTGGTATCTCTGGTCTATGGTTTACCAGAAAGAGCTAAAGGAGATGACGACAA 349
Db 255 CAATGGATTTGGTATCTCTGGTCTATGGTTTACCAGAAAGAACTGAGAGATGACGACAA 314
Qy 350 GTACCGGCTGCACCTGTCTAGTGGCTGACCTCTCTTTGTCTATCACACTCCCCCTTTGGGC 409
Db 315 GTACAGGCTGCACCTGTCTAGTGGCGCACCTCTCTTTGTCTATCACACTCCCCCTTTGGGC 374
Qy 410 AGTTGATGCCATGGCTGACTGTGTTGTTGGAAATTTTGTGTAAGSCTGTCTCATATCAT 469
Db 375 AGTTGATGCCATGGCTGACTGTGTTGTTGGAAATTTTGTGTAAGSCTGTCTCATATCAT 434
Qy 470 CTACACTGTCAACCTCTTACAGCAGCGCTTCTCATCTCTGGCTTTCATCAGGCTTGACCGGTA 539
Db 435 CTACACAGTCAACCTCTTACAGCAGCTGTCTCTCTCTGGCTTTCATCAGTCTGACCGCTA 494
Qy 530 CTTGCGCATTTGTCACGCGCACCAAGCTCAAGGCGCAAGAACTGCTGGCTGAAAGGC 589
Db 495 CTTGCGCATTTGTCACGCGCACCAAGCTCAAGGCGCAAGAACTGCTGGCTGAAAGGC 554
Qy 590 AGTCTATGTGGGCTGTGGATCCAGCCCTCTCTCTGACTATATACCTGACTTCTCTTTC 649
Db 555 GGTCTATGTGGGCTGTGGATCCCTGCTCTCTGCTGACTATTTCCCGACTTCTCTTTC 614
Qy 650 CGAGCTCAGCCAGGGGGACATCATAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTA 709
Db 615 CAACG-----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTCTA 659
Qy 710 CCCGATAGCTGTGGATGGTGTGTTCAATCCAGCATATATATGGTGGGTCTCATCT 769
Db 660 CCCCATGACTTGTGGGTGGTGTGTTCCAGTTTCAGCACATCATGGTGGCTTATCT 719
Qy 770 GCGCGGATCGTCACT 829
Db 720 GCTGTGTTATGTCATCT 779
Qy 830 GGGCCACCAAGGCGAAGGCGCTCAAGAGCAGTCATCTCTCTCTCTCTCTCTCTCTCTCT 889
Db 780 GGGCCACCAAGGCGAAGGCGCTCAAGAGCAGTCATCTCTCTCTCTCTCTCTCTCTCTCT 839
Qy 890 CTGCTGCTGCCATATTTATGTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCTCTCT 949
Db 840 CTGTTGGCTGCCCTTACTTACATTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCTCT 899
Qy 950 CAAGCAGGATGTGACTTTCAGAGCATTTGTCAGAGTGGATCTCCATCAGAGGCGCT 1009
Db 900 CAAGCAGGATGTGACTTTCAGAGCATTTGTCAGAGTGGATCTCCATCAGAGGCGCT 959
Qy 1010 CGCCTTCT 1069
Db 960 AGCTTTCTTCCACTGTGTCTGAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
Qy 1070 AAGCTCTGCCAGCATGCACTCAACTCAATGAGAGAGGCTCCAGGCTCAGAGTCTCTTTC 1129
Db 1020 AACTCTCTGCCAGCATGCACTCAACTCAATGAGAGAGGCTCCAGGCTCAGAGTCTCTTTC 1079
Qy 1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCAGGAGTCAAGATCTCTCTCTCTCTCT 1189
Db 1080 CAAAGGAAAGCGGGGTGGACACTCTTCCGTCTCCAGGAGTCAAGATCTCTCTCTCTCT 1139
Qy 1190 CTCAGGTAACCCCTTATGCAAGACTTATATATATATATATATATATATATATATATAT 1249
Db 1140 CTCAGGTAACCCCTTATGCAAGACTTATATATATATATATATATATATATATATATAT 1191

Db 660 CCCAATGACTTGTGGTGGTGTGTCAGTTTCAGCACATCATGTTGGCTTATCCT 719
Qy 770 GCCCGGATGTCATCCTCTCGTTTACTGCAATCATCTCTAAGCTGTCACTCCAA 829
Db 720 GCCTGTATTTGTCATCCTCTCGTTTACTGCAATCATCTCTAAGCTGTCACTCCAA 779
Qy 830 GGCCACCAAGAGCGAGCCCTCAAGACGAGTCATCTCATCTAGCTTTCTTTGC 889
Db 780 GGCCACCAAGAGCGAGCCCTCAAGACGAGTCATCTCATCTAGCTTTCTTTGC 839
Qy 890 CTCTGGCTGCCATATATGTGGGATCAGATCGACTCTCTATCTTTTGGAGTCA 949
Db 840 CTCTGGCTGCCATATATGTGGGATCAGATCGACTCTCTATCTTTTGGAGTCA 899
Qy 950 CAGCAGAGATGACTTCCAGAGCATGTGCAGAGTGCATCTCATCAGAGGCGCT 1009
Db 900 CAGCAGAGATGACTTCCAGAGCATGTGCAGAGTGCATCTCATCAGAGGCGCT 959
Qy 1010 CGCTTTCTTCCACTGTGCTGAGAGCCCTCTCTATGCTTCTCGGGGCGAAGTTCAA 1069
Db 960 AGCTTTCTTCCACTGTGCTGAGAGCCCTCTCTATGCTTCTCGGGGCGAAGTTAA 1019
Qy 1070 AGCTTTCTTCCACTGTGCTGAGAGCCCTCTCTATGCTTCTCGGGGCGAAGTTTC 1129
Db 1020 AGCTTTCTTCCACTGTGCTGAGAGCCCTCTCTATGCTTCTCGGGGCGAAGTTTC 1079
Qy 1130 CAAAGGAAAGCGGGTGGACACTCTCTCGTCTCCAGGAGTCAGATCTCTCAGTTTCA 1189
Db 1080 CAAAGGAAAGCGGGTGGACACTCTCTCGTCTCCAGGAGTCAGATCTCTCAGTTTCA 1139
Qy 1190 CTCAGTAAACCTTATGCAAGACTTATATATATATATATATATATATATATAGAACTT 1249
Db 1140 CTCAGTAAACCTTATGCAAGACTTATATATATATATATATATATATATATAGAACTT 1191
Qy 1250 TTTTATGTTACATTTTCAGATATAAGAGTGCACCTCTGTGACAGTTTCTTTTCTTTT 1309
Db 1192 TTTTATGTTACATTTTCAGATATAAGAGTGCACCTCTGTGACAGTTTCTTTTCTTTT 1251
Qy 1310 TTTTATGTTACATTTTCAGATATAAGAGTGCACCTCTGTGACAGTTTCTTTTCTTTT 1369
Db 1252 TTTTATGTTACATTTTCAGATATAAGAGTGCACCTCTGTGACAGTTTCTTTTCTTTT 1305
Qy 1370 AT 1429
Db 1306 TTTTATA-----TAAATTTTTTTTGTGTTTATATATATATATATATATATATATAT 1361
Qy 1430 AAGTTCTTAGTAGTGTATCTGTGTAGGACTGTAGAACTGTAGAGGAAAGAACTGA 1489
Db 1362 AAGTTCTTAGTAGTGTATCTGTGTAGGACTGTAGAA-----AAGGAACTGA 1413
Qy 1490 ACATTCAGAAATGTGGTAAATGAATAAGCTAGCCGTGATCTCTCAGCTGTGCTGA 1549
Db 1414 ACATTCAGAAATGTGGTAAATGAATAAGCTAGCCGTGATCTCTCAGCTGTGCTGA 1473
Qy 1550 TAATCTCTTCAATCCGAGGAGCAGCCCAACCCCAACCCCAACCCCAACCCCAACCTTCTTAAT 1609
Db 1474 TAGATAATCTCTCA-----1488
Qy 1610 TGTGTGTTATGCTGTGATGATGTTGTTGTTTGTGTTTGTGTTGTTGTTGTTTGT 1669
Db 1489 -----TTCCGTTGGAACGTTTTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1525
Qy 1670 TTTTCTCTAAAGATGCACTTAAACCAAGAGCTGAATATGTTGTTGTTGTTGTTGTTGTTGTT 1729
Db 1526 TTTTCTCTAAAGATGCACTTAAACCAAGAGCTGAATATGTTGTTGTTGTTGTTGTTGTTGTT 1581
Qy 1730 TTTTCTCTAAAGATGCACTTAAACCAAGAGCTGAATATGTTGTTGTTGTTGTTGTTGTTGTT 1789
Db 1582 -----TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1623
Qy 1790 GTACAGCTTTGTTAT 1849
Db 1624 GTACAGCTTTGTTAT 1683

Qy 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 5

US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,478
; FILING DATE: 25-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-805-478-3

Query Match 51.9%; Score 974.4; DB 2; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

Qy 50 GCAGTAGACAGTACCCCTCTCAGGCGTTTGGTGTCCGGTAACCCACCGCTGTAGAGC 109
Db 20 GCGCGCGCAAGTACGCGCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCTGAGAAC 79
Qy 110 GAGTGTTCCTTGAACCGATCAGTGTGATATATACACTTCTGATTAACCTCTGAAGA 169
Db 80 CAGCGGTTTACCATGA-----GGGATCAGTATATACACTTTCAGATAACTACACCGAGGA 134
Qy 170 AGTGGGCTCGAGACTATGACTCCAAAGGAACCTGCTTCGGGATGAAAAACGTCCA 229
Db 135 AATGGGCTCAGGGAGACTATGACTCCATGAGGAACCTGTTTCCGTGAAGAAATGCTAA 194

QY	230	TTTCAATAGGATCTTCTGTGCCCACTCTACTTTCATCATCTTCTTGACTGGCATAGTCGG	289
Db	195	TTTCAATAAAATCTTCTGTGCCCACTACTCTCCATCATCTTCTTAACTGGCATTTGTGG	254
QY	290	CAATGGATTGGTCATCTCTGTGTCAATGGTTTACCAGAGAGAGCTAAGAGAGCATGACGACAA	349
Db	255	CAATGGATTGGTCATCTCTGTGTCAATGGTTTACCAGAGAGAACTGAGAAGCANTGACGACAA	314
QY	350	GTACCGGCTGCACCTGTGCAGTGGCTGACCTCCCTCTTTGTTCATCACACTCCCTTCTGGGC	409
Db	315	GTACAGGCTGCACCTGTGCAGTGGCGGACCTCTTTGTTCATCACGCTTCCCTTCTGGGC	374
QY	410	AGTTGATGCCATGGCTGACTGGTACTTTTGGGAAATTTTGTGTGAAGCTGTCCATATCAT	469
Db	375	AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACTTCTCTATGCAAGGCGAGTCCATGTCA	434
QY	470	CTACACTGTCAAACCTCTTACAGCAGCGTTCTCATCTCTGGCGCTTCATCAGCGCTGACCGGTA	529
Db	435	CTACACAGTCAAACCTCTACAGCAGTGTCTCATCTCTGGCGCTTCATCAGTCTGACCGCTA	494
QY	530	CTTGGCCATTGTCCACGCCACCAACAGTCAAAGGCCAAGGAACTGTGTGGCTGAAGAAGC	589
Db	495	CCTGGCCATCGTCCACGCCACCAACAGTCAAGAGGCCAAGGAACTGTGTGGCTGAAGAAGT	554
QY	590	AGTCTATGTGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCATCTCTTGC	649
Db	555	GGTCTATGTGGCGTCTGGATCCCTGCGCTCTCTGTGACTATTCCGACATTCATCTTGC	614
QY	650	CGACGTCAAGCCAGGGGACATCACTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTA	709
Db	615	CAACG-----TCAGTAGGCAGATGACAGATATATCTGTGACCGCTTCTA	659
QY	710	CCCGGATAGCCTGTGGATGGTGTGTTTCAATTCACGATATATATGGTGGGTCTCATCTCT	769
Db	660	CCCAATAGCTGTGGGTGGTGTGTGTCCAGTTTCAGCACATCATGGTGGCCTTATCCT	719
QY	770	GCCCGCATCGTCATCTCTCTGTGTACTGCATCATCATCTCTAAGCTGTCAACATCCAA	829
Db	720	GCCTGGTATTGTTCATCTCTCTCTGCTATTGCAATTATCATCTCCAAGCTGTCAACATCCAA	779
QY	830	GGGCCACCAAGAGGGGAGGCGCTCAAGACGACAGTCATCTCATCTCTAGCTTCTTTGC	889
Db	780	GGGSCACAGAAGCGCAAGCGCTCAAGACACAGTCATCTCATCTCTGGCTTCTTTCGC	839
QY	890	CTGCTGCTGCCATATTATGTGGGATCAGCATCGACTCTCTCATCTCTTTGGAGTTCAT	949
Db	840	CTGTTGGCTGCTTACTACATTGGGATCAGCATCGACTCTCTCATCTCTCTGGAATTCAT	899
QY	950	CAAGCAAGGATGTGACTTCGAGAGCAATTGTGCACAAGTGGATCTCCATCAAGAGCCCT	1009
Db	900	CAAGCAAGGATGTGAGTTTGAGATTTCAGAACATGTGTGCAACAGTGGATTCCATCACCGAGGCCCT	959
QY	1010	CGCTTCTTCCACTGTTTCCGTGACCCCATCTCTATGCCTTCTCTGGGCGCAAGTCAA	1069
Db	960	AGCTTTCTTCCACTGTTTGTGTGAACCCCATCTCTATGCTTCTCTGGAGCCAAATTTAA	1019
QY	1070	AAGCTCTGCCAGCATCGACTCAACTCCATGAGCAGAGGCTCCAGCCCTCAAGATCTCTTTC	1129
Db	1020	AACCTCTGCCAGCAGCACTCACTCTGTGTGAGCAGAGGCTCCAGCCCTCAAGATCTCTCTC	1079
QY	1130	CAAAGGAAAGGGGGTGGACACTCTTCCGTCTCCACGAGTCAGAATCTCTCCAGTTTCA	1189
Db	1080	CAAAGGAAAGCGAGTGGACATTCACTGTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCA	1139
QY	1190	CTCCAGCTAACCCCTTATGCAAGACTTTATATAATATATATATATATATATATATATATAT	1249
Db	1140	CTCCAGCTAA-----CACAGATGTAAAGACTTTTTTTTATACGATAAATAACTTTT	1199
QY	1250	TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTTGTACAGTTTTTTTTT	1309
Db	1192	TTTTAAGTTACACATTTTTCAGATATAAAGACTGACCAATTTTGTACAGTTTTTATTCG	1251

Qy	1310	TTTTTAATTCAGCTGTGGGAGTTATATGTCCTCTAGTTTTTGTGAGGTTTGACTTAATTT	1369
Db	1252	TTGTTGGAATTTTGT-----CTTGTGTTCTTTTAGTTTTTGTGAAGTTTAATTGACATTA	1305
Qy	1370	ATATAAATATGTTTTTTTGTGTTTCACTGTAAGAGCGCTCTAGGACGACCTGTGGCC	1429
Db	1306	TTTATA-----TAAATTTTTTTTGTGTTTCATATTGATGTGCTTAGGACGACCTGTGGCC	1361
Qy	1430	AAGTCTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTGAACCTGTAGAGGAAGAACTGA	1489
Db	1362	AAGTCTCTTAGTCTGATGTCTCGTGTAAGGACTGTAGAA-----AAGGGAAGCTGA	1413
Qy	1490	ACATTCAGAAATGTGTGTGTAATAATGAATAAAGCTAGCCGTGATCTCAGCTGTGCTGCA	1549
Db	1414	ACATTCAGAGCGTGTAGTGAATACGTAAGCTAGAAATGATCCGACGTGTTTATGCA	1473
Qy	1550	TAATCTCTTCATTTCCGAGGAGCACCCACCCACCCACCCACCCACCCCATTCCTTAAAT	1609
Db	1474	TAGATAATCTCTCCA-----TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1488
Qy	1610	TGTTTGGTTATGCTGTGTGATGGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTTTT	1669
Db	1489	-----TTCCCGTGGAAAGTCTGTTTCTGTTCTTTAAGACGTGAT	1525
Qy	1670	TTTTCTGTAAGATGCGACTTAAACCAAAGCTGAATGGTGGTAGAAAAGCTCGGGGT	1729
Db	1526	TTTGTGTGAAGATGGCACTTATACCAAAGCCAAAGTGGT-ATAGAAATGCTGG---	1581
Qy	1730	TTTTTTTTTTGTTGTTTTTTTTCAGTTTTCAGAGTAGATTGACTTCAGTCCCTACAAAT	1789
Db	1582	-----TTTTTCAGTTTTTCAGAGTGGGTTGATTTTCAGACCTAC-AGT	1623
Qy	1790	GTACAGTCTTTGATTACATTGTTTAAATAAAGTCAATGATAACTTAAAAAATAAAAAA	1849
Db	1624	GTACAGTCTTTGATTAAAGTGTGTTAATAAAGTACAAGTTAACTTAAAAAATAAAAAA	1683
Qy	1850	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	1877
Db	1684	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	1711

RESULT 6

US-08-802-627A-3
Sequence 3, Application US/0802627A
Patent No. 5892017
GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,627A
FILING DATE: 19-Feb-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-802-627A-3

Query Match      51.9%; Score 974.4; DB 2; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

QY 50 GCAGGTAGCAGTACCCCTCTGAGCGCTTGGTCTCGGTAAACCACACCGGCTGTAGAC 109
DB 20 GCAGCGGCAAGTGAAGCGCGAGCGGCTGAGTCTCCAGTAGCCACCGCATCTGGAGAAC 79
QY 110 GAGTGTGCCATGGAACCGATCAGTGTGAGTATATACATCTCTGATAAACTACTCTGAAGA 169
DB 80 CAGCGGTACCATGGA-----GGGGATCAGTATATACATCTCAGATAAATACACCGAGA 134
QY 170 AGTGGGCTGTGGAGATGATGACTCCAAAGAAAGCCCTGCTCCGGGATGAAAGCGTCCA 229
DB 135 AATGGGCTCAGGGGACATGATGACTCCATGAAGAAAGCCCTGCTCCGGTGAAGAAAGTCTAA 194
QY 230 TTTCATATGAGATCTCTCTGCCACCATCTACTTTCATCTCTCTGACATGATGCGG 289
DB 195 TTTCATATAAATCTTCTGCCACCATCTACTTTCATCTCTCTGACATGATGCGG 254
QY 290 CAATGATGGTATGATCTGTGTATGGTTTACCAAGAAAGCTTAAGAGATGACGAGACAA 349
DB 255 CAATGATGGTATGATCTGTGTATGGTTTACCAAGAAAGCTTAAGAGATGACGAGACAA 314
QY 350 GTACCGGCTGACCTGTGAGTGGCTGACCTCTCTTTGTCATCACACTCCCTCTGGGC 409
DB 315 GTACAGGCTGACCTGTGAGTGGCTGACCTCTCTTTGTCATCACACTCCCTCTGGGC 374
QY 410 AGTTGATGCCATGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
DB 375 AGTTGATGCCATGGCTGACTGTGACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCGTCTCATCTCTGGCTTCATCAGCTGACCGGTA 529
DB 435 CTACAGCTCAACCTCTACAGCAGCGTCTCATCTCTGGCTTCATCAGCTGACCGGTA 494
QY 530 CTTGGCCATTTGCCAGCCACCAACAGTCAAGGCAAGAAAGCTCTGGCTGAAAGGC 589
DB 495 CTTGGCCATTTGCCAGCCACCAACAGTCAAGGCAAGAAAGCTCTGGCTGAAAGGC 554
QY 590 AGTCTATGTGGGCTGTGATGCCAGCTCTCTCTGATCTACCTGATCTTCACTGATCTTTC 649
DB 555 GGTCTATGTGGGCTGTGATGCCAGCTCTCTCTGATCTTCACTGATCTTTC 614
QY 650 CGAGCTCAGCGAGGAGACATCAGTCAAGGCGGATGACAGTACATCTGTGACCGCTTTA 709
DB 615 CAACG-----TCAGTGAGGAGATGACAGTATATCTGTGACCGCTTTA 659
QY 710 CCGCATAGCTGTGATGGTGTGTTTCAATTCAGCATATAATGGTGGGCTCATCCT 769
DB 660 CCGCAATGACTGTGGGCTGTGTTTCCAGTTTCAGCATCATGTTGGGCTTATCCT 719
QY 770 GCCGGCATGCTCATCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTGACATCCNA 829
DB 720 GCCTGGTATGTCTCTCTCTGTTACTGTCATCATCTCTTAAGCTGTGACATCCNA 779
```

```

QY 830 GGGCCACCAAGCGCGAGCCCTCAAGACGACAGTCACTCTCTATCTCTAGCTTCTTTGC 889
DB 780 GGGCCACCAAGCGCGAGCCCTCAAGACGACAGTCACTCTCTATCTCTAGCTTCTTTGC 839
QY 890 CTGCTGGCTGCATATATTATGTTGGGATCAGCATCGACTCTTCTATCTCTTTGGAGTCAT 949
DB 840 CTGTTGGCTGCCTTACTACATTGGGATCAGCATCGACTCTTCTATCTCTCTGGAATCAT 899
QY 950 CAAGCAAGGATGTGACTTTCGAGAGCATGTGCAAGTGGATCTCCATCTCAGAGGCGCT 1009
DB 900 CAAGCAAGGCTGTGAGTTTGAGAAACACTGTGCAAGTGGATTTCCATCTCAGAGGCGCT 959
QY 1010 CGCCTTCTTCCACTGTTCCCTGAAACCCATCTCTATGCTCTCTCGGSCCAAGTTCAA 1069
DB 960 AGCTTTCTTCCACTGTTCTGAAACCCATCTCTATGCTCTCTCTCGAGCCAAATTTAA 1019
QY 1070 AAGCTCTGCCAGCATCACTCAACTCCATGAGCAGAGGTCCTCAGCTCAAGATCTTTTC 1129
DB 1020 AACCTCTGCCAGCAGCAGCTCACTCTGTGAGCAGAGGTCCTCAGCTCAAGATCTCTC 1079
QY 1130 CAAGGAAAGCGGCTGAGACACTCTTCCGCTCTCCAGGATCAGAACTCTCCAGTTTCA 1189
DB 1080 CAAGGAAAGCGGCTGAGACACTCTTCCGCTCTCCAGGATCAGAACTCTCCAGTTTCA 1139
QY 1190 CTCAGGCTAACCCCTTATGCAAGACTTATATAATATATATATATATATATGATAAGAACTT 1249
DB 1140 CTCAGGCTAA-----CACAGATGTAAAGACTTTTTTTTATACGATAAATAACTTTT 1191
QY 1250 TTTTATGTTACACATTTTCCAGATATAAGAGACTGACCACTCTTGTACAGTTTTTTTT 1309
DB 1192 TTTTAAAGTTACACATTTTCCAGATATAAGAGACTGACCAATATTTGTACAGTTTATATGC 1251
QY 1310 TTTTAAATTTGACTCTTGGGAGTTTATGTTCTCTAGTTTGTGAGGTTTGACTTAATTT 1369
DB 1252 TTGTTGGATTTTGT-----CTTGTGTTCTTTAGTTTGTGAACTTAAATGACTTA 1305
QY 1370 ATATAAATATGTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1429
DB 1306 TTTATA-----TAAATTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1361
QY 1430 AAGTTCTTAGTGTGTTTATCTGTGTGAGGACTGTGAGAACTGTAGAGAAAGAACTGA 1489
DB 1362 AAGTTCTTAGTGTGTTTATCTGTGTGAGGACTGTGAGAA-----AAGGGAACCTGA 1413
QY 1490 ACATTCAGAAATGTGTTGTTAAATGAAATGAGCTGAGCGTGTGCTCAGCTGTGTTGCTGA 1549
DB 1414 ACATTCAGAGCGTGTAGTGAATCAGTAAAGCTAGAAATGATCCCAAGCTGTTTATGCA 1473
QY 1550 TAATCTCTTCAATTCGAGGAGCACCCCAACCCCAACCCCAACCCCAACCCCAATCTTAAAT 1609
DB 1474 TAGATAATCTCTCCA-----TTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1488
QY 1610 TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1669
DB 1489 -----TTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1525
QY 1670 TTTTCTGTAAAGATGCGACTTTAAACCAAGCCTGAAATGGTGGTGAATGCTGGGT 1729
DB 1526 TTTCTGTAGAGATGCGACTTTAAACCAAGCCTGAAATGGTGGTGAATGCTGGGT 1581
QY 1730 TTTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1789
DB 1582 -----TTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1623
QY 1790 GTAAGCTCTGTATTAATTTAAAGTCAATGATTAAGTCAATTAAGTCAATTAAGTCAATTAAGT 1849
DB 1624 GTAAGCTCTGTATTAATTTAAAGTCAATGATTAAGTCAATTAAGTCAATTAAGTCAATTAAGT 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
DB 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711
```


[illegible]

APPLICANT: Lee, James
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: PF4A Receptors
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPattn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/03/104,296
 FILING DATE: 24-June-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/701265
 FILING DATE: 22-AUG-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/664228
 FILING DATE: 06-JUN-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/076093
 FILING DATE: 11-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/810782
 FILING DATE: 19-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Love, Richard B.
 REGISTRATION NUMBER: 34,659
 REFERENCE/DOCKET NUMBER: P0706P2C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-5530
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1737 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear

US-09-104-296--3

	Query Match	51.9%; Score 974.4; DB 3; Length 1737;
	Best Local Similarity	76.1%; Pred. No. 2.3e-195;
	Matches 1391; Conservative	0; Mismatches 301; Indels 136; Gaps 10;
Qy	50	GCAGGTAGCAGTACCCCTCTGAGGCGCTTTGGTCTCCGGTTAACACCAACCGCTGTAGAC 109
Db	20	GC CGCGCGCAAAAGTAGCGCGGAGGCGCTGAGTCTCCAGTAGGCCACCGCATCTGAGAAC 79
Qy	110	GAGTGTTCCTACGGAAACCGATCAGTGTGAGTATATACACTTTTGATAAATCTACTCTGAAGA 169
Db	80	CACGGTTACCATGGA-----GGGGATCAGTATATACATTTAGATAAATCTACACCGAGGA 134
Qy	170	AGTGGGGTCTGGAGACTATGACTCCAAAGGAACCCCTGCTCCGGGATGAAGAACGTCCA 229
Db	135	AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTTCCGTGAAGAAAATGCTAA 194
Qy	230	TTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCTCTTCTTGACTGGCATACTCGG 289
Db	195	TTTCAATAAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTCTTAACGGCATTTGGGG 254
Qy	290	CAATGATTGGTGATCTCTGTGTCTATGGTTTACCAGAAGAAGCTAAGGAGCATGACGGACAA 349
Db	255	CAATGAATTGGTTCATCTCTGTGTATGGTTTACCAGAAGAAGCATGAGAAGCATGACGGACAA 314
Qy	350	GTACCGGCTGCACCTGTCTAGTGGCTGACCTCTCTTTTGTCATCACACTCCCTTTCTGGGC 409

RESULT 9
US-09-104-236-3
; Sequence 3, Application US/09104296
; Patent No. 6087475
; GENERAL INFORMATION:

STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-06380-2

Query Match 51.9%; Score 974.4; DB 5; Length 1737;
Best Local Similarity 76.1%; Pred. No. 2.3e-195;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;
QY 50 GCAGGTAGCAGTGACCCCTCTGAGGCGTTTGGTGTCCGGTAACACACACCGCTGTAGAGC 109
DB 20 GCGCGGCGCAAGTCAGCGCGAGGCGCTGAGTGTCTCAGTAGCCACCGCATCTGGAGAC 79
QY 110 GAGTGTGCGATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
DB 80 CAGCGGTTACCATGGA-----GGGATCAGTATATACACTTCAAGATACTACACCGAGA 134
QY 170 AGTGGGCTCGAGACTATGACTCCAAACAGGACCCCTGCTCCGGATGAAGAGTCCA 229
DB 135 AATGGGCTCAGGGGACTATGACTCCATGAGGACCCCTGTTCCGTGAAGAAATGCTAA 194
QY 230 TTTCATAGGATCTTCTCGCCACCATCTACTTCATCATCTTCTTGACTGGCATAGTCGG 289
DB 195 TTTCATAAATCTTCTCGCCACCATCTACTCCATCATCTCTTAACCTGGCAITGGG 254
QY 290 CAATGGATGTGATCTCTGCTCATGGTTACAGAGAACTAAGAGCATGACGCAAA 349
DB 255 CAATGGATGTGATCTCTGCTCATGGTTACAGAGAACTAAGAGCATGACGCAAA 314
QY 350 GTACCGGCTGACCTGTGAGTGGCTGAGCTCTCTCTTGTCTATCATCACTCCCTCTGGC 409
DB 315 GTACAGGCTGACCTGTGAGTGGCGGACCTCTCTTGTCTATCAGCTTCTCTCTGGC 374
QY 410 AGTTGATGCGATGGCTGATGCTGACTTTTGGGAAATTTTGTGAAGGTGTCATATCAT 469
DB 375 AGTTGATGCGGTGGGAACTGCTACTTTTGGGAACTTCTCTATGCAAGGCACTGCTATCAT 434
QY 470 CTACACTGTCAACCTCTACAGCAGGCTTCTCATCTGCGCTTCTATCAGCTGGACCGGTA 529
DB 435 CTACACAGTCAACCTCTACAGCAGTGTCTCTATCTGCGCTTCTATCAGTCTGGACCGCTA 494
QY 530 CTTCCGCAATGTCACGCGCAACACAGTCAAGGCAAGGAACTGCTGGCTGAAAAGGC 589
DB 495 CTTGGCCATGTCACGCGCAACACAGTCAAGGCAAGGAACTGTTGGCTGAAAAGGT 554
QY 590 AGTCTATGTGGCGTCTGATCCAGCCCTCTCTGACTATATCTGACTTCACTTCTTTC 649
DB 555 GGTCTATGTGGCGTCTGATCCCTCTCTGACTATATCTGACTATCTCCGACTTCACTTTC 614
QY 650 CGACGTACCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTA 709
DB 615 CAACG-----TCAGTGAGCAGATGACAGATATATCTGTGACCGCTTCTA 659
QY 710 CCCGATAGCCTGTGGATGGTGGTGTTCATATTCAGCATATATGTTGGTCTCTATCCT 769
DB 660 CCCCATGACTTGTGGGTGGTGTGTTCAGTTCAGCACATCATGTTGGGCTTATCCT 719
QY 770 GCGCGCATGTCATCTCTCTCTGATCATCATCTATCTAAGTGTGTCACACTCAA 829
DB 720 GCTGTGATGTGATCTCTCTCTGATGATGATATCATCTCAAGTGTGTCACACTCAA 779
QY 830 GGGCCACAGAGCGGAGCCCTCAAGACAGATCATCTCTATCTCTAGTCTTCTTTC 889
DB 780 GGGCCACAGAGCGGAGCCCTCAAGACAGATCATCTCTATCTCTAGTCTTCTTTC 839
QY 890 CTGCTGGCTGCTATATATGTTGGGATCAGATGATCTCTCTCTCTCTTGGAGTCA 949
DB 840 CTGTTGGCTGCTTACTACTATGTTGGGATCAGATGATCTCTCTCTCTCTTGGAAATCAT 899
QY 950 CAAGCAAGGATGTGACTTTCAGAGCATGTGTCACAAAGTGGATCTCCATCATCAGAGGCCCT 1009
DB 900 CAAGCAAGGATGTGACTTTCAGAGCATGTGTCACAAAGTGGATTTCCATCATCAGAGGCCCT 959
QY 1010 GCGCTTCTTCCACTGTTGCTGAAACCCCATCTCTATGCTCTCTCGGGGCGCAAGTTCAA 1069

DB 960 AGCTTCTTCCACTGTTGTCTGAACCCCATCTCTATGCTTCTTCTGGAGCCAAATTTAA 1019
QY 1070 AAGCTTGGCCAGCATGCACTCACTTCCATGAGCAGAGGCTCCAGCTCAAGATCTCTTTC 1129
DB 1020 AAGCTTGGCCAGCAGCATCTACCTCTGTGAGCAGAGGCTCCAGCTCAAGATCTCTCTC 1079
QY 1130 CAAAGGAAAGCGGGTGGACACTCTTCCGTCTCCAGGAGTCAAGATCTCTCCAGTCTTCA 1189
DB 1080 CAAAGGAAAGCGAGTGGACATCTCTCTGTTCCACTGAGTCTGAGTCTTCAAGTCTTCA 1139
QY 1190 CTCAGCTAACCTTATGCAAGACTTATATATATATATATATATATATATATATATATAT 1249
DB 1140 CTCAGCTAA-----CACAGATGTAAGAGACTTTTTTTTATACGATAAATAAATTTT 1191
QY 1250 TTTTATGTTACACATTTTCCAGATATAAGAGACTGACAGCTCTGACAGGACCTGTGGCC 1309
DB 1192 TTTTAAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGACAGTCTTTTATGTC 1251
QY 1310 TTTTAAATGACTGTTGGGAGTATTATGTTCTCTAGTCTTCTTGTGAGGTTTGACTTAATTT 1369
DB 1252 TTCTTGAATTTTGT-----CTTGTGTTTCTTGTAGTCTTGTGAAAGTTTAAATTGACTTA 1305
QY 1370 ATATAAATATTGTTTGTGTTTGTGTTTTCATGTGATGAGCTGTAGCAGGACCTGTGGCC 1429
DB 1306 TTTATA-----TAAATTTTGTGTTTCTTGTATGATGTTGTCTAGCAGGACCTGTGGCC 1361
QY 1430 AAGTCTTCTAGTACTGTTTATCTGTGTGAGACTGTAGAACTGTAGAGGAAAGAACTGA 1489
DB 1362 AAGTCTTCTAGTCTGATGTTCTGTTGTTAGGACTGTAGAA-----AAGGAACTGA 1413
QY 1490 ACATTCAGAAATGTGTGTAATTAAGTAAAGTACCCCTGATCTCTCAGCTGTGTGCTGCA 1549
DB 1414 ACATTCAGAGGCTGTGATGTAATCAGTAAAGTAGAAATGATCCCGAGCTGTGTTATGCA 1473
QY 1550 TAATCTCTTCTATCTCCAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1609
DB 1474 TAGATAATCTCTCCA----- 1488
QY 1610 TGTGTTGTTATGCTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1669
DB 1489 -----TCCCGTGGAAAGTCTTCTCTGTTCTTAAGACGTGAT 1525
QY 1670 TTTTCTGTAAGAGTGGCACTTAAACCAAGCCTGAATGTTGTTAGAAATGCTGGGCT 1729
DB 1526 TTTGCTGTAGAGATGGCACTTAAACCAAGCCTGAATGTTGTTAGAAATGCTGG----- 1581
QY 1730 TTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1789
DB 1582 -----TTTTTCAGTCTTTCAGGAGTGGGTTGATTTTCAGCACCTAC-AGT 1623
QY 1790 GTACAGTCTTGTATACATGTTTAAAGTCAATGATAAAGTCAATGATAAAGTCAATGATAA 1849
DB 1624 GTACAGTCTTGTATGATGTTTAAAGTCAATGATAAAGTCAATGATAAAGTCAATGATAA 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
DB 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 11

US-09-517-605-14
; Sequence 14, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Inso
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605

QY 45 CAGGTGAGGTAGTACAGTACCTCTGAGGCGTTTGGTGTCCGGTAAACACACCGCTGT 104
Db 1 CGGCAGCAGTACGAAAGTACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTGG 60
QY 105 AGAGCGAGTGTGCCATGAAACGATGAGTGTGAGTATATACACTTCTGATACACTCTCT 164
Db 61 AGAACAGCGGTACCATCGA-----GGGGATCAGTATATACACTTCTGATACACTACACC 115
QY 165 GAAGAAGTGGGTCTGGAGACTATGACTCCAAAGGAACCTGCTCCGGGATGAAAC 224
Db 116 GAGGAATGGCTCAGGAGACTATGACTCCAAAGGAACCTGCTCCGGTGAAGAAAT 175
QY 225 GTCCATTTCAATAGATCTCTGCCACCATCTACTCATCATCTCTTCTGAGTGGGATA 284
Db 176 GCTAAATTTCAATAAATCTCTGCCACCATCTACTCCATCATCTCTTAACTGGCAAT 235
QY 285 GTCCGCAATGGATGTGTGATCTGTGTATCCAGGAAGCTTAAGGAGCATGACG 344
Db 236 GTGGGCAATGGATGTGTGATCTGTGTATCCAGGAAGCTTAAGGAGCATGACG 295
QY 345 GACAAAGTCCGCTGCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 404
Db 296 GACAAAGTCCGCTGCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 355
QY 405 TGGGCAATGGATGTGTGATCTGTGTATCCAGGAAGCTTAAGGAGCATGACG 464
Db 356 TGGGCAATGGATGTGTGATCTGTGTATCCAGGAAGCTTAAGGAGCATGACG 415
QY 465 ATCATCTACACTGTCAACCTCTACAGCAGCTTCTCATCTCCGCTTCTCATCAGCTGGAC 524
Db 416 GTCATCTACAGTCAACCTCTACAGCAGCTTCTCATCTCCGCTTCTCATCAGCTGGAC 475
QY 525 CGGTACTCGCATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 584
Db 476 CGGTACTCGCATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 535
QY 585 AAGGCAAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 644
Db 536 AAGTGTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 595
QY 645 TTGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 704
Db 596 TTGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 640
QY 705 CTTTACCCGATAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 764
Db 641 TTTACCCGATAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 700
QY 765 ATCTGCGCGCATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 824
Db 701 ATCTGCGCGCATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 760
QY 825 TCCAGGGGCGCAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 884
Db 761 TCCAGGGGCGCAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 820
QY 885 TTGCGCTGTGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 944
Db 821 TTGCGCTGTGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 880
QY 945 GTCATCAAGCAAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1004
Db 881 ATCATCAAGCAAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 940
QY 1005 GCGCTGCGCTTCTTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1064
Db 941 GCGCTGCGCTTCTTCCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1000
QY 1065 TTCAAAGCTCTGCCAGCATGCTCAACTCCATGAGCAGAGGCTCCAGCTCTCAAGATC 1124
Db 1001 TTTAAACCTCTGCCAGCATGCTCAACTCTGTGAGCAGAGGCTCCAGCTCTCAAGATC 1060
QY 1125 CTTTCAAAGCAAGCGGGGTGGACACTCTTCTGCTTCCAGGAGTCAGATCTCTCCAGT 1184

Db 1061 CTCTCCAAAGGAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAGT 1120
QY 1185 TTTCACTCCAGCTAAACCCCTTATGCAAGACTTATATAATATATATATATATATATATATATAT 1244
Db 1121 TTTCACTCCAGCTAAACAGATGATAAAGACTTTTT-----TTTATACGATAAATAA 1172
QY 1245 AACTTTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTGTACAGTTTTT 1304
Db 1173 CTTTTTTTTAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTTGTACAGTTTTT 1232
QY 1305 TTTTTTTTTAATGACTGTGGAGTTTATGTTTCTCTAGTTTTTTTGTGAGGTTTGCATT 1364
Db 1233 ATTGCTGTGTTGGAATTTTGT-----CTTGTGTTTCTTGTAGTTTTTGTGAAGTTTAAATG 1286
QY 1365 AATTATATAAATATGTTTTTTTGTGTTTCTGTAATGAGCGTCTAGGCGAGGACCTG 1424
Db 1287 ACTATTTATA-----TAAATTTTTTTTGTGTTTCTGTAATGAGCGTCTAGGCGAGGACCTG 1342
QY 1425 TGCCAAAGTCTTTAGTAGCTGTTTATCTGTGTGAGACTGTGAGAACTGTAGAGGAAGAA 1484
Db 1343 TGCCAAAGTCTTTAGTGTGTATGTCGTGTGAGACTGTAGAA-----AAGGGA 1394
QY 1485 ACTGAACATTCAGAAATGTGTGTAATGAAATAAGCTAGCGGTGATCTCAGCTGTGTTG 1544
Db 1395 ACTGAACATTCAGAGCGTGTGTTAATCAGTAAAGCTAGAAATGATCCCCAGCTGTGTT 1454
QY 1545 CTGCATA 1551
Db 1455 ATGCATA 1461

RESULT 13

US-09-023-655-1213
; Sequence 1213, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhauer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 1213:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1664 base pairs

[illegible]

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0002 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 1235:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1225 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GENBANK
 CLONE: g189313
 US-09-016-434-1235

Query Match 45.8%; Score 859; DB 4; Length 1225;
 Best Local Similarity 83.3%; Pred. No. 3.3e-171;
 Matches 1009; Conservative 0; Mismatches 180; Indels 22; Gaps 2;

QY 136 TGAGTATATACACTCTGATTAACCTCTGAAGAAGTGGGTCTGGAGACTATGACTCCA 135
 DB 36 TCAGTATATACACTCTGATTAACCTCTGAAGAAGTGGGTCTGGAGACTATGACTCCA 95

QY 196 ACAGGAACCCCTGCTCCGGGATGAAGAGCTGCATTTCAATAGGATCTTCTGCCACCA 255
 DB 96 TGAAGGAACCCCTGCTCCGGGATGAAGAGCTGCATTTCAATAGGATCTTCTGCCACCA 155

QY 256 TCTACTTCATCATCTCTTGTACTGTCATGTCGCAATGGATTTGGTATCTGTCATGG 315
 DB 156 TCTACTTCATCATCTCTTGTACTGTCATGTCGCAATGGATTTGGTATCTGTCATGG 215

QY 316 GTTACCAAGAAGCTAAGGAGCATAGCGCAAGTACCGGCTCCACTGTCAGTGGCTG 375
 DB 216 GTTACCAAGAAGCTAAGGAGCATAGCGCAAGTACCGGCTCCACTGTCAGTGGCTG 275

QY 376 ACCTCCTCTTTGTATCATCATCTCCCTCTCTGGGAGTTGATGCCATGGCTGACT 435
 DB 276 ACCTCCTCTTTGTATCATCATCTCCCTCTCTGGGAGTTGATGCCATGGCTGACT 335

QY 436 TTGGGAATTTTGTGTAAAGCTGTCCATATCATCTACATGTCAAACTCTACAGCAGCG 495
 DB 336 TTGGGAATTTTGTGTAAAGCTGTCCATATCATCTACATGTCAAACTCTACAGCAGCG 395

QY 496 TTCTCATCTGGCTTTCATAGCTTGAACCGGTACCTGCCATTTGTCCACGCCACCA 555
 DB 396 TCCTCATCTGGCTTTCATAGCTTGAACCGGTACCTGCCATTTGTCCACGCCACCA 455

QY 556 GTCRAAGGCCAAGGAATCTGCTGGCTGAAAGGAGTCTATGTGGGGTCTGGATCCCG 615
 DB 456 GTCAGAGGCCAAGGAATCTGCTGGCTGAAAGGAGTCTATGTGGGGTCTGGATCCCG 515

QY 616 CCCTCCTCCTGACTATACCTGACTTTCATCTTTGGCGAGCTCAGCCAGGGGAGCATGCTG 675
 DB 516 CCCTCCTCCTGACTATACCTGACTTTCATCTTTGGCGAGCTCAGCCAGGGGAGCATGCTG 560

QY 676 AGGGGGATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGTGGTGT 735
 DB 561 AGGCAGATGACAGATATATCTGTGACCGCTTTACCCCGATAGCTGTGGATGTGGTGT 620

QY 736 TTCAATTCAGCATATAATGTGGGTCTCATCTGCGCGGATCGTCACTCTCTCTGTT 795
 DB 621 TCCAGTTTCAGCATATGTTGGTCTTATCTGCTGTTATTTGTCATCTCTGCTGCT 680

QY 796 ACTGATCATCTCTAAGCTGTACATCTCAAGGGCCACAGAGCGGAGGCGCTCA 855
 DB 681 ATTGATTTATCTCTCAAGCTGTACATCTCAAGGGCCACAGAGCGGAGGCGCTCA 740

QY 856 AGACACAGTCACTCTCATCTCTAGTTCTTTGCTGCTGCTGCTGCTATTTATGCTGGGA 915

DB 741 AGACCACAGTCATCTCATCTCTGCTGCTTCTTCTGCTGCTGCTGCTTACTACATGGGA 800
 QY 916 TCAGCATCGACTCCTCTCTATCTCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGACA 975
 DB 801 TCAGCATCGACTCCTCTCTATCTCTCTGGAATTCATCAAGCAAGGATGTGAGTTGAGACA 860
 QY 976 TTGTGCACAGTGGATCTCCATCAGAGGCGCTTCTTCCACTGTTGCTGCTGAACC 1035
 DB 861 CTGTGCACAGTGGATTTCCATCAGGAGGCGCTTCTTCCACTGTTGCTGCTGAACC 920
 QY 1036 CCATCTCTATGCTTCTCTCGGGGCAAGTTCAAAAGCTCTGCCAGCATGCACTCAACT 1095
 DB 921 CCATCTCTATGCTTCTCTGGAGCCAAATTTAAACCTCTGCCAGCATGCACTCAACT 980
 QY 1096 CCATGAGCAGAGCTCCAGCTCAAGATCTTTCAAAGGAAGCGGGTGGACACTCTT 1155
 DB 981 CTGTGACAGAGGCTCCAGCTCAAGATCTCTCCAAAGGAAGCGGGTGGACACTTCAAT 1040
 QY 1156 CGGTCTCCAGGAGTCAGATCTCTCAGTTTCTCAGCTTAAAGCTTATGCAAGACT 1215
 DB 1041 CTGTTTCCACTGAGTCTGAGTCTTCAAGTTTCTCCAGCTTAAACAGATGTAAA--- 1096
 QY 1216 TATATAATATATATATATATATGATAAAGAACTTTTTTATGTTACATTTTCCAGATAT 1275
 DB 1097 ---AGACTTTTTTTTATACGATAAATRACTTTTTTTAAAGTTACACATTTTTCAGATAT 1153
 QY 1276 AAGAGACTGACAGCTCTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTAT 1335
 DB 1154 AAAAGACTGACCAATATGTACAGTTTTTATGCTGTTGGATTTTGTCTGTGTTCT 1213
 QY 1336 GTTCTCTTAGT 1346
 DB 1214 TTAGTTTTGT 1224

RESULT 15

US-08-153-848-45
 ; Sequence 45, Application US/08153848
 ; Patent No. 5759804
 ; GENERAL INFORMATION:
 ; APPLICANT: Godiska, Ronald
 ; APPLICANT: Gray, Patrick W.
 ; APPLICANT: Schweikart, Vicki L.
 ; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
 ; NUMBER OF SEQUENCE: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Bicknell,
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/153,848
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/977,452
 ; FILING DATE: 17-NOV-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5759804and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 31794
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 1317 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 201..1211

US-08-153-848-45

Query Match 44.5%; Score 834.6; DB 1; Length 1317;

Best Local Similarity 83.9%; Pred. No. 4.5e-166;

Matches 977; Conservative 0; Mismatches 164; Indels 24; Gaps 2;

```
QY 134 TGTGAGTATATACACTCTCTGATTAACACTACTCTGAGAAAGTGGGCTCGGAGACTATGACTC 193
DB 164 TTGTCAGATATACACTTTCAGATTAACACTACACCGAGGAAATGGGCTCAGGGGACTATGACTC 223
QY 194 CAACAAGGAACCTGTCTTCGGGATGAAACGTCCTCAATTTCAATAGGATCTTCTGCCCCAC 253
DB 224 CATGAGGAACTCTGTTCCGTGAAGAAATGCTAATTTCAATAAATCTTCTGCCAC 283
QY 254 CATCTACTTCATCATCTTCTTGATGCTGGCATAGTCGGCAATGATTTGGTATCCTGGTCAAT 313
DB 284 CATCTACTCCATCATCTTCTTAACCTGGCATTTGGGCAATGGATTGGTCACTGGTCAAT 343
QY 314 GGGTTACCAAGAAGCTAGGAGCATGACGGAAGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
DB 344 GGGTTACCAAGAAGCTAGGAGCATGACGGAAGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
QY 374 TGACCTCTCTCTTTGTATCATCACTCCCTTTCTGGGAGTTGATGCCATGGCTGACTGGTA 433
DB 404 CGACCTCTCTTTGTATCATCACTCCCTTTCTGGGAGTTGATGCCATGGCTGACTGGTA 463
QY 434 CTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACACTGTCTCAACTCTACAGAG 493
DB 464 CTTTGGGAATTTTGTGTAAGCTGTCCATATCATCTACACTGTCTCAACTCTACAGAG 523
QY 494 CGTTCTCATCTCGGCTTTTCATAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 553
DB 524 TGTCCTCATCTCGGCTTTTCATAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 583
QY 554 CAGTCAAGGCCAAGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
DB 584 CAGTCAAGGCCAAGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
QY 614 AGCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
DB 644 TGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688
QY 674 TCAGGGGATGACAGGTATATCTGACCGCTTTTACCGGATAGCTGCTGCTGCTGCTGCTGCTGCT 733
DB 689 TGAGGCAGATGACAGATATATCTGACCGCTTTTACCGGATAGCTGCTGCTGCTGCTGCTGCTGCT 748
QY 734 GTTTCATATCCAGCATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
DB 749 GTTTCAGTTTCAGCATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 808
QY 794 TTACTGCATCATCTCTAAGCTGTACACTCCAAAGGGCCACCAAGCGCAAGGCCCT 853
DB 809 CTATTGCAATTATCATCTCCAAAGCTGTACACTCCAAAGGGCCACCAAGCGCAAGGCCCT 868
QY 854 CAAGACGACAGTATCTCTCATCTCTAGCTTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913
DB 869 CAAGACGACAGTATCTCTCATCTCTAGCTTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
QY 914 GATCAGCATCGACTCTCTCTCATCTCTTGGAGTCAATCAAGCAGGATGTGACTTCGAGAG 973
DB 929 GATCAGCATCGACTCTCTCTCATCTCTTGGAGTCAATCAAGCAGGATGTGACTTCGAGAG 988
QY 974 CATTGTGACAAAGTGGATTCATCAAGAGGCCCTCGGCTTCTTCCACTGTTGCTGAA 1033
```

Search completed: May 17, 2004, 16:15:39
Job time : 171.281 secs

```
DB 989 CACTGTGCACAAAGTGGATTTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGTCTGAA 1048
QY 1034 CCCATCCTCTATGCCCTTCTCGGGCCAAAGTTCAAAAAGCTCTGCCAGCATGCACCTCAA 1093
DB 1049 CCCATCCTCTATGCTTTCTTCTTGGAGCCAAATTTAAAACCTCTGCCAGCAGCCTCAC 1108
QY 1094 CTCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAGCGGGTGGACACTC 1153
DB 1109 CTCTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAGCGAGTGGACATTC 1168
QY 1154 TTCGCTCTCCACGGAGTCAGAATCCTCCAGTTTTCTCCAGCTTAACCTTTATGCAAGA 1213
DB 1169 ATCTGTTTTCCACTGAGTCTGAGTCTTCAAGTTTTCTCCAGCTTAACACAGATGTA 1226
QY 1214 CTTATATAATATATATATATATATATATATATATATATATATATATATATATATATAT 1273
DB 1227 -----AGACTTTTTTATACGATAAATACCTTTTTTTTAAAGTTACACATTTTTCAGAT 1279
QY 1274 ATAAAGAGCTGACCAAGTCTTTGTACA 1298
DB 1280 ATAAAGAGCTGACCAATATTGAAAA 1304
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 12:11:00 ; Search time 918.766 Seconds

(without alignments)
9271.041 Million cell updates/sec

Title: US-09-367-052-1

Perfect score: 1877

Sequence: 1 ccactcctaatacgactcact.....aaaaaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5094648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	974.4	51.9	1737	9	US-09-104-063-3		Sequence 3, Appli
2	974.4	51.9	1737	13	US-10-666-689-3		Sequence 3, Appli
3	952.8	50.8	1679	13	US-10-211-462-80		Sequence 80, Appli
4	952.8	50.8	1679	13	US-10-181-906-9		Sequence 9, Appli
5	952.8	50.8	1679	13	US-10-342-887-912		Sequence 912, App
6	952.8	50.8	1679	13	US-10-151-274-14		Sequence 14, Appli
7	952.8	50.8	1679	13	US-10-172-118-912		Sequence 912, App
8	952.8	50.8	1679	13	US-10-170-385-332		Sequence 332, App
9	952.8	50.8	1679	15	US-10-225-567A-75		Sequence 75, Appli
10	952.8	50.8	1679	15	US-10-021-660-58		Sequence 58, Appli
11	952.8	50.8	1679	15	US-10-341-434-177		Sequence 177, App
12	952.8	50.8	1711	10	US-09-971-392-20		Sequence 20, Appli
13	952.8	50.8	1711	10	US-10-101-510-459		Sequence 459, App
14	951.8	50.7	1664	17	US-10-641-643-1213		Sequence 1213, App

Sequence 2143, Ap
Sequence 636, App
Sequence 1, Appli
Sequence 77, Appli
Sequence 741, App
Sequence 1235, App
Sequence 674, App
Sequence 84, Appli
Sequence 125, App
Sequence 2, Appli
Sequence 143, App
Sequence 143, App
Sequence 484, App
Sequence 1, Appli
Sequence 789, App
Sequence 687, App
Sequence 14955, A
Sequence 14955, A
Sequence 2611, Ap
Sequence 2611, Ap
Sequence 8936, Ap
Sequence 8936, Ap
Sequence 8936, Ap
Sequence 162, App
Sequence 65, Appli
Sequence 5, Appli

15 944.8 50.3 1670 9 US-09-880-107-2143
16 944.8 50.3 1670 10 US-09-960-706-636
17 944.8 50.3 1670 16 US-10-372-683-1
18 944.8 50.3 1670 16 US-10-440-464-77
19 934.4 49.8 5161 16 US-10-160-401-1
20 859 45.8 1225 15 US-10-101-510-741
21 859 45.8 1225 16 US-10-305-720-1235
22 847 45.1 1224 15 US-10-101-510-674
23 809.6 43.1 1059 13 US-09-813-651B-84
24 809.6 43.1 1059 15 US-10-014-322A-125
25 809.6 43.1 1059 16 US-10-014-322A-125
26 808 43.0 1102 9 US-09-870-759-143
27 808 43.0 1102 10 US-09-751-708A-143
28 804.8 42.9 1059 11 US-09-826-509-484
29 576.4 30.7 1902 9 US-09-953-692-1
30 576.4 30.7 1902 9 US-09-953-717-1
31 530 28.2 1076 15 US-10-017-161-789
32 530 28.2 1076 16 US-10-292-798-687
33 345.2 18.4 507 13 US-10-085-783A-14955
34 345.2 18.4 507 16 US-10-242-535A-14955
35 303.2 16.2 421 9 US-09-796-693-2611
36 303.2 16.2 421 15 US-10-040-862-2611
37 303.2 16.2 421 16 US-10-057-475B-2611
38 303.2 16.2 421 16 US-10-154-884B-2611
39 287 15.3 487 9 US-09-796-692-8936
40 287 15.3 487 15 US-10-040-862-8936
41 287 15.3 487 16 US-10-057-475B-8936
42 287 15.3 487 16 US-10-154-884B-8936
43 198.4 10.6 1119 12 US-09-997-722-162
44 198.4 10.6 1119 15 US-10-251-385-65
45 198.4 10.6 1679 9 US-09-104-063-5

ALIGNMENTS

RESULT 1
US-09-104-063-3
; Sequence 3, Application US/09104063
; Patent No. US20020168356A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winstatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,063
; FILING DATE: 24-June-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/701265
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA: 08/664228
; APPLICATION NUMBER: 08/664228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5930
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-104-063-3

Query Match      51.9%; Score 974.4; DB 9; Length 1737;
Best Local Similarity 76.1%; Pred. No. 5.3e-214;
Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;

QY 50 GCAGGTAGCAGTACCCCTCTGAGGGGTTTGGTCCGGTACACCAACCGGCTGTAGAGC 109
Db
QY 110 GAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGTAACTACTCTGAAGA 169
Db 80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGA 134
QY 170 AGTGGGGTCTGAGACTATGACTCCAAAGGAACCCCTGCTCCGGGATGAAGAGTCCA 229
Db 135 AATGGGCTCAGGGGACTATGACTCCATGAAGAACCCCTGTTCCGTTGAAGAAATGCTAA 194
QY 230 TTTCATAGAGTCTTCCTCCCAACCATCTACTATCATCTCTTCTGACTGGCATAGTCGG 289
Db 195 TTTCATATAAACTCTCTCCCAACCATCTACTCTCATCTCTTCTTAATGCAITGTGGG 254
QY 290 CAATGATTTGGTATCTCTGTCATGGTTACAGAAAGCTTAAGAGCATGAACGACAA 349
Db 255 CAATGATTTGGTATCTCTGTCATGGTTACAGAAAGCTTAAGAGCATGAACGACAA 314
QY 350 GTACCGGCTGCACCTGTCTAGTGGCTGACCTCTCTTTGTTCATCACACTCCCTCTCTGGGC 409
Db 315 GTACAGGCTGCACCTGTCTAGTGGCGGCTCTCTCTTTGTTCATCACGCTCTCCCTCTGGGC 374
QY 410 AGTTGATGCAATGGCTGACTGTCTCTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
Db 375 AGTTGATGCGGTGGCAACTGTCTCTTTGGGAAATTTCTATGCAAGGCGAGTCCATGTCT 434
QY 470 CTACACTGTCAACCTCTACAGCAGGCTTCTCATCTCTGGCTTTCATCAGCCTGGACCGGTA 529
Db 435 CTACACAGTCAACCTCTACAGCAGGCTCTCTCATCTCTGGCTTTCATCAGTCTGGACCGGTA 494
QY 530 CTTCCGCAATGTCCACGCCCAACAGTCAAGGCCAAGGAACTGTGGCTGAAAGGC 589
Db 495 CTTGGCCATCTGTCCACGCCCAACAGTCAAGGCCAAGGAACTGTGGCTGAAAGGT 554
QY 590 AGTCTATGTGGGCTGTGATCCAGCCCTCTCTGCTACTATACCTGACTTCACTCTTTCG 649
Db 555 GGTCTATGTGGGCTGTGATCCCTGCTCTCTGCTGACTATTCCTCGACTTCACTCTTTCG 614
QY 650 CGACGTCAGCCAGGGGACATCAGTCAGGGGATGACAGGTATCATCTGTGACCGGCTTTTA 709
Db 615 CAAGC-----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTTA 659
QY 710 CCCCAGTACGCTGTGATGGTGGTGTTCATTTCAATTCAGGCATATATTAATGGTGGTCTCATCT 769
Db 660 CCCCAGTACGCTGTGATGGTGGTGTGTGTTCCAGTTTCAGCACATCATGGTGGCTTATCTCT 719
QY 770 GCCCGGCATCTGCTCTCTCTGTGTACTGTCATCATCTCTTAAGCTGTGCACCTCCAA 829
Db 720 GCTGTGATTTGTCATCTCTCTCTGTCTATTTGCAATATCATCTCCAAGCTGTGCACCTCCAA 779
QY 830 GGGCCACCAAGAGCGGACGCCCTCAAGACGACAGTCACTCATCTCTAGCTTTCTTTGC 889

```

RESULT 2

```

Db 780 GGGCCACCAAGAGCGCAAGGCCCTCAAGACCACAGTCACTCTCATCTGCTTTCTTCGC 839
QY 890 CTGCTGGCTGCCATATTTATGTGGGGATCAGCATCGACTCCTCTCATCTTTTGGAGTCAT 949
Db 840 CTGTGGCTGCCCTTACTACATTTGGGATCAGCATCGACTCCTCTCATCTCTGCGAAATCAT 899
QY 950 CAAGCAGGATGTGACTTCGAGAGCATTTGTCACAGTGTGATCTCCATCAGAGCCCT 1009
Db 900 CAAGCAGGATGTGAGTTTGAAGAACATCTGTGCAAGTGTGATTTCCATCAGGAGCCCT 959
QY 1010 CGCTTCTTCCACTGTTCCTGAACCCCATCTCTATGCTTCTCTCGGGCCCAAGTTCAA 1069
Db 960 AGCTTCTTCCACTGTTCCTGAACCCCATCTCTATGCTTCTCTTGGAGCCAAATTTAA 1019
QY 1070 AAGCTCGCCAGCATGCACTCAACTCCATGAGCAGAGCTCCAGCTCAAGTCTCTTTC 1129
Db 1020 AAGCTCGCCAGCATGCACTCAACTCTGTGAGCAGAGGTCAGGCTCAAGATCCTCTC 1079
QY 1130 CAAAGGAAAGCGGGGTGGACACTCTTCCGCTCTCCAGGAGTCAAGATCTCTCCAGTTTCA 1189
Db 1080 CAAAGGAAAGCGAGGTGGACATTCATCTGTTCCACTGAGTCTGAGTCTTCAAGTTTCA 1139
QY 1190 CTCAGCTAACCTTATGCAAGACTTATATATATATATATATATATATATATATATAT 1249
Db 1140 CTCAGCTAA-----CACAGATGTAAGAGACTTTTTTTTATACGATAAATAACTTTT 1191
QY 1250 TTTTATGTTACACATTTTCCAGATATAAGAGACTGACAGCTGTGTACAGTTTCTTTT 1309
Db 1192 TTTTATGTTACACATTTTCCAGATATAAGAGCTGACCAATATTGTACAGTTTCTTATGC 1251
QY 1310 TTTTAAATGACTGTGGGAGTTTATGTTCTCTAGTTTGTGAGTTTGTGACTTAATTT 1369
Db 1252 TTGTGGATTTTGT-----CTTGTGTTTCTTTAGTTTGTGAAAGTTTAATTGACTTA 1305
QY 1370 ATATAATATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1429
Db 1306 TTTATA-----TAAATTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1361
QY 1430 AAGTTCTTAGTACTGTTTATCTGTGTGAGGACTGTAGAACTGTAGAGGAAAGAACTGA 1489
Db 1362 AAGTTCTTAGTCTGTATGCTCTCGTGTAGGACTGTAGAA-----AAGGAACTGA 1413
QY 1490 ACATTCAGAAATGTGGTAAATTAAGTAGCCGCTGATCCTCAGCTGTGCTGCA 1549
Db 1414 ACATTCAGAGCGTGTAGTGAATCACGTAAGCTAGAAATGATCCCCAGCTGTTTATGCA 1473
QY 1550 TAATCTCTTCAATCCGAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCAT 1609
Db 1474 TAGATAATCTCTCCA----- 1488
QY 1610 TGTGTGTTATGCTGTGATGGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1669
Db 1489 -----TTCCCGTGGAAACGTTTTCCTGTTCTTAAGACGTGAT 1525
QY 1670 TTTTCTGTAAGATGGCACTTAAACCAAGCCTGAATTTGGTGTAGAAATGCTGGGGT 1729
Db 1526 TTTGCTGTAGAAATGGCACTTAAACCAAGCCTGAATTTGGTGTAGAAATGCTGG--- 1581
QY 1730 TTTTGTGTTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1789
Db 1582 -----TTTTTCAGTTTTCAGGAGTGGGTTGATTTTCAGCACCCTAC-AGT 1623
QY 1790 GTACAGTCTTGTATTTACATTTGTTAAATAAGTCAATGATAAAGCTTAAAGGAGGAGG 1849
Db 1624 GTACAGTCTTGTATTTAGTTTGTGTTAAATAAGTACATGTTAAAGGAGGAGGAGG 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

```

US-10-666-689-3			US-10-666-689-3		
; Sequence 3, Application US/10666689			Query Match		
; Publication No. US20040037830A1			51.98; Score 974.4; DB 13; Length 1737;		
; GENERAL INFORMATION:			Best Local Similarity 76.1%; Pred. No. 5.3e-214;		
; APPLICANT: Lee, William I.			Matches 1391; Conservative 0; Mismatches 301; Indels 136; Gaps 10;		
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and					
; FILE REFERENCE: P0706P2C2D2C1					
; CURRENT APPLICATION NUMBER: US/10/666,689					
; PRIOR FILING DATE: 2003-09-19					
; PRIOR APPLICATION NUMBER: US 09/104,063					
; PRIOR FILING DATE: 1988-06-24					
; PRIOR APPLICATION NUMBER: US 08/701,265					
; PRIOR FILING DATE: 1996-08-22					
; PRIOR APPLICATION NUMBER: US 08/664,228					
; PRIOR FILING DATE: 1996-06-06					
; PRIOR APPLICATION NUMBER: US 08/076,093					
; PRIOR FILING DATE: 1993-06-11					
; PRIOR APPLICATION NUMBER: US 07/810,782					
; PRIOR FILING DATE: 1991-12-19					
; PRIOR APPLICATION NUMBER: US 07/677,211					
; PRIOR FILING DATE: 1991-03-29					
; NUMBER OF SEQ ID NOS: 6					
; SEQ ID NO 3					
; LENGTH: 1737					
; TYPE: DNA					
; ORGANISM: Homo sapien					
US-10-666-689-3					
QY	50	GCAGGTAGCAGTACCTCTGAGCGCTTGGTGTCCGGTAACACACACCGCTGTAGAGC	109	CGAGTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGATCCTTTC	1129
DB	20	GCGCGGCGCAAGTGACCGCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCTGGAGAC	79	AACTCTGCCAGCAGCGCACTCACCTCTGTGAGCAGAGGGTCCAGCTCAAGATCCTCTC	1079
QY	110	GAGTGTGGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATTAATCTCTGAAGA	169	CAAGGAAAGCGGGTGGACACTCTTCCTGTCCACGAGAGTCAGTAATCCTCCAGTTTTCA	1189
DB	80	CAGCGGTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGA	134	CAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTGTGAGTCTTCAAGTTTTCA	1139
QY	170	AGTGGGTCTGGAGACTATGACTCCACAGGACCGCTGCTCCGGATCAAAACGTCCTCA	229	CTCCAGCTTAAACCTTATGCAAGACTTATAATAATATATATATATATATATATATATAT	1249
DB	135	AATGGGTCTAGGGAGTATGACTCCATGAGGAACCTGTTTCCGTGAGAAATGCTAA	194	CTCCAGCTTAA-----CACAGATGTAAGAGACTTTTTTTTATACGATAAATAACTTTT	1191
QY	230	TTTCAATAGGATCTTCTGCGCCACCATCTACTTTCATCATCTCTTGTGACTGGCATGCG	289	TTTTTAAATGACTGTGGGAGTTTATGTTCTCTAGTTTCTTGTGAGTTTGTGACTTAAATTT	1369
DB	195	TTTCAATAAATCTTCTGCGCCACCATCTACTTTCATCATCTCTTGTGACTGGCATGCG	254	TTGTTGGATTTTGT-----CTTGTGTTTCTTTAGTTTCTTTAGTTTAAATGACTTA	1305
QY	290	CAATGGATGTGATCTCTGCTGATGTTTACAGAGAGCTAGGAGCATGACGGACAA	349	ATATAAATATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	1429
DB	255	CAATGGATGTGATCTCTGCTGATGTTTACAGAGAGCTAGGAGCATGACGGACAA	314	TTTATA-----TAAATTTTTTTTTTTTTCATATTGATGTGTCTAGGCAGGACCTGTGGCC	1361
QY	350	GTACGGCTGCACCTGTGACGAGGCTGACCTCTCTTTGTATCATCACTCCCTCTCTGGGC	409	AACTTCTTAGTAGCTGTTTATCTGTGTGAGACTGTAGACTGTAGAGGAGAACTGA	1489
DB	315	GTACAGGCTGCACCTGTGACGAGGCTGACCTCTCTTTGTATCATCACTCCCTCTCTGGGC	374	AACTTCTTAGTAGCTGTTTATCTGTGTGAGACTGTAGACTGTAGAGGAGAACTGA	1413
QY	410	AGTTGATGCCATGCTGATCTGTTTGGAAATTTTGTGAAAGTGTGAGGCTGCTCATATCAT	469	ACATTCACAGAAAGTGTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	1549
DB	375	AGTTGATGCCATGCTGATCTGTTTGGAAATTTTGTGAAAGTGTGAGGCTGCTCATATCAT	434	ACATTCACAGAAAGTGTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	1473
QY	470	CTACACTGTCACTCTACAGAGGCTTCTCATCTGCGCTTCTCATGAGCTGGAGCCGTA	529	TAACTCTTTTCTATCCGAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCTTAAT	1609
DB	435	CTACACTGTCACTCTACAGAGGCTTCTCATCTGCGCTTCTCATGAGCTGGAGCCGTA	494	TAGATAATCTCTCCA-----	1488
QY	530	CCTCGCCATTTGCCAGGCCACCAACAGTCAAGGCCCAAGGAACTGTGGCTGAAAGGC	589	TGTTTGGTTATGCTGTGTGATGGTTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGT	1669
DB	495	CCTGGCCATCTGTCACGCCACCAACAGTCAAGGCCCAAGGAACTGTGGCTGAAAGGC	554	-----TTCCCGTGAAACGTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1525
QY	590	AGTCTATGTGGCGTCTGATATCCAGCCCTCTCTCTGATATATCTGATCTCATCTTTGC	649	TTTTTCTGTAAGAGATGGCACTTAAACCAAGCCTGAAATGGTGGTAGAATGCTGGGT	1729
DB	555	GGTCTATGTGGCGTCTGATATCCAGCCCTCTCTCTGATATATCTGATCTCATCTTTGC	614	TTTGCTGTAGAGATGGCACTTAAACCAAGCCTGAAAGTGGT-ATAGAAATGCTGG- - -	1581

QY 1730 TTTTTCCTGTTGTTTGTGTTTTCAGTTTTCAGAGTAGATTGACTTTCAGTCCCTACAAAT 1789
Db 1582 -----TTTTTCAGTTTTCAGGAGTGGTTGATTTTCAGCACCTAC-AGT 1623
QY 1790 GTACAGTCTTCTGATTACATTTTAAATAAAAGTCAATGATTAACATTAATAAAAAA 1849
Db 1624 GTACAGTCTTCTGATTAAAGTTTAAATAAAAGTCAATGTTAACTTAATAAAAAA 1683
QY 1850 AAAAAAAAAAAAAAAAAAAAAAAAAA 1877
Db 1684 AAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 3
US-10-211-462-80
; Sequence 80, Application US/10211462
; Publication No. US20040033495M1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-462-80

Query Match 50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

QY 44 GCAGGTGCAGGTAGCAGTCAACCTCTGAGGCGTTTGGTGTCTCCGGTAACCAACCGGCTG 103
Db 12 GCGGCAGCAGGTAGCAAGTGAAGCGCGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTACTC 163
Db 72 GAGAACCCAGCGGTACCATGGA-----GGGGATCAGTATATACACTTCCAGATACTAC 126
QY 164 TGAAGAGTGGGTCTGGAGATATACATCCACAGGAACCTGCTCCGGGATGAAAA 223
Db 127 CGAGGAATGGGCTCAGGGGACTATGATCCATGAGGAACCTGTTCCGTTGAAGAAA 186
QY 224 CGTCCATTTCAATPAGGATCTTCTGCCACCACTACTACTTCACTTCTTCTGAGTGGCAT 283
Db 187 TGCTAATTTCAATAAATCTTCTGCCACCACTACTACTCATCTTCTTAACTGGCAT 246
QY 284 AGTCGCAATGGATTGGTGTGATCTGCTGATGGTGTACAGGAAGACTAAGGAGCATGAC 343
Db 247 TGTGGGCAATGGATTGGTGTGATCTGCTGATGGTGTACAGGAAGACTGAGGAAGCATGAC 306
QY 344 GGACAAAGTACCGGCTGCACCTGTGAGTGGTGAACCTCTTGTGATACACACTCCCTTT 403
Db 307 GGACAAAGTACAGGCTGCACCTGTGAGTGGTGGACCTCTTGTGATCACGCTTCCCTTT 366
QY 404 CTGGGCGAGTTGATGCCATGCGCTGACTGTGACTTGTGGAAATTTTGTGTAAGGCTGTCCA 463

Db 367 CTGGCAGTTGATGCGGTGGCAAACTGGTACTTTTGGGAACCTCTCTATGCAAGGAGTCCA 426
QY 464 TATCATCTACACTCTCAACCTCTACAGCAGCGTCTCATCTCTGGCTTTCATCAGCCTGGA 523
Db 427 TGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTCTCATCTGGCTTTCATCAGTCTGA 486
QY 524 CCGGTACCTTCGCTGATTTGTCAGCCCAACACAGTCAAAAGGCCAAAGAACTGCTGGCTGA 583
Db 487 CCGGTACCTGGCCATCTGTCACCGCCACCAACAGTCAAGAGGCCAAAGAACTGTTGGCTGA 546
QY 584 AAAGGAGTCTATGTGGGGTCTGGATCCAGCCCTCTCTGATATACCTGATACCTGACTTCA 643
Db 547 AAAGGTGGTCTATGTGGGGTCTGGATCCCTCTCTGCTGACTATTTCCGACTTTCAT 606
QY 644 CTTTGGCCAGCTCAGCCAGGGGACATCAGTCAAGGGGATGACAGTACATCTGTGACCG 703
Db 607 CTTTGGCCAGC-----TCAGTGAGGAGATGACAGATATATCTGTGACCG 651
QY 704 CTTTACCCCGATAGCTGTGATGGTGTGTTTCAATTCAGCATATAATAGTGGTCT 763
Db 652 CTTTACCCCAATGACTTGTGGGTGGTGTGTTTCCAGTTTCAGCACATCATGTTGGCT 711
QY 764 CATCTGCGCGCATGCTCATCTCTCTGTTTACTGTGATCATCATCTCTAAAGCTGTGACA 823
Db 712 TATCTGCTCTGTTGTGTCATCTCTGCTGCTATTGCAATTAATCATCTCCAGCTGTACA 771
QY 824 CTCGAAGGCCACCAAGAGCGCAAGCCCTCAAGAGGACAGTCACTCTCTAGCTTT 883
Db 772 CTCGAAGGCCACCAAGAGCGCAAGCCCTCAAGACCACAGTCACTCTCTCTGCTTT 831
QY 884 CTTTGTCTGCTGCTGCCATATTATGTGGGGATCAGCATCGACTCTCTCTCTCTTTGGG 943
Db 832 CTTTGTCTGTTGGTCTGCTTACTACATTGGGATCAGCATCGACTCTCTCTCTCTGGA 891
QY 944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTCACAAAGTGGATCTCCATCACAGA 1003
Db 892 AATCATCAAGCAAGGATGTGAGTTTGAGAACACTGTGTCACAAAGTGGATTTCCATCACCGA 951
QY 1004 GGGCTCTGCTTCTTCCAGTGTGCTGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
Db 952 GGGCTCTGCTTCTTCCAGTGTGCTGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
QY 1064 GTTCAAAAGCTCTGCGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGAT 1123
Db 1012 ATTTAAACCTCTGCGCCAGCAGCACTCACTCTGTGAGCAGAGGCTCCAGCTCAAGAT 1071
QY 1124 CTTTCCAAAGGAAAGGGGGTGGACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
Db 1072 CTTCTCCAAAGGAAAGGGGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG 1131
QY 1184 TTTTCACTCCAGCTAACCTTATGCAAGACTTTATATAATATAATATAATATAATATAATA 1243
Db 1132 TTTTCACTCCAGCTAACAGATGTAAGACTTTTT-----TTTATACGATAAATA 1183
QY 1244 GAACCTTTTATGTTACATATTTCCAGATATAAGAGACTGACCACTCTTGTACAGTTTT 1303
Db 1184 ACTTTTTTTTAAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTGTACAGTTTT 1243
QY 1304 TTTTCTTTTAAATGACTCTGCGAGTTTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1363
Db 1244 TATTGCTGTGTTGATTTTGT-----CTTGTGTTCTTTTGTGTTTGTGAGTTTAAAT 1297
QY 1364 TAAATTTATAAATATTGTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1423
Db 1298 GACTTATTATA-----TAAATTTTTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 1353
QY 1424 GTGGCAAGTTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGAGCTGTAGAGGAAGA 1483
Db 1354 GTGGCAAGTTCTTAGTTGCTGTATGTTCTGTGTGAGTGTAGAA-----AAGGG 1405
QY 1484 AACTGAACATTCACAGATGTGTGTAATTAATGAATAAAGCTAGCCGCTGATCTCTCAGCTGTT 1543

Db 1406 AACTGAACATTCAGAGCGGTAGTGAATCACTAAAGCTAGAAATGATCCCGACGTGT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473

RESULT 4
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karsenty, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducey, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181.906
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match 50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

Qy 44 GCAGGTGAGGTAGCAGTGCCTCTCAGGCGTTTGGTGCTCGGTAAACACACACGCGTG 103
Db 12 GCGCAGCAGGTAGCAAGTGAACCCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCG 71
Qy 104 TAGAGCGAGTGTCCCATGGAAACCGATCAGTGTGAGTATATACACTTCTCATCTCTC 163
Db 72 GAGAACACGCGGTACCATGGA-----GGGATCAGTATATACACTTCCAGTAACATAC 126
Qy 164 TGAAGAGTGGGCTGAGAGTATGACTCCAAAGGAAACCTGCTCCGCGATGAAA 223
Db 127 CGAGAAATGGGCTCAGGGGACTATGACTCCATGAGGAAACCTGTTCCGTGAGAAA 186
Qy 224 CGTCCATTTCAATAGGATCTTCTGCCACCATCTACTTCTCATCTCTTCTGATGGCAT 283
Db 187 TGCTAAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTAAGTGGCAT 246
Qy 284 AGTGGCAATGATGTGATCCTGGTATGGTTACGAGAGAGAGTAAAGGAGCATGAC 343
Db 247 TGTGGCAATGATGTGATCCTGGTATGGTTACGAGAGAGAGTAAAGGAGCATGAC 306
Qy 344 GGACAAGTACCGGCTGCACCTGTGAGTGGCTGACCTCTCTTGTCTATCAGCTCCCTTT 403
Db 307 GGACAAGTACCGGCTGCACCTGTGAGTGGCTGACCTCTCTTGTCTATCAGCTCCCTTT 366
Qy 404 CTGGCAGTGTGATGCGATGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 463
Db 367 CTGGCAGTGTGATGCGTGGCAAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 426
Qy 464 TATCATCTACAGTGTCAACCTCTACAGAGCGTTTCTCATCTGCGCTTCTCATCGCTGGA 523
Db 427 TGTCTATCTACAGTGTCAACCTCTACAGAGCGTTTCTCATCTGCGCTTCTCATCGCTGGA 486
Qy 524 CCGGTACCTCGCATTTGTCAGGCAACCAAGTCAAAAGGCAAGGAACTGCTGGCTGA 583
Db 487 CGGCTACCTGGCATCTGTCAGGCAACCAAGTCAAAAGGCAAGGAACTGCTGGCTGA 546
Qy 584 AAAGCAGTCTATGTGGGCTGCTGATCCAGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 643
Db 547 AAAGGTGGTCTATGTGGGCTGCTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606

RESULT 5
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong

Qy 644 CTTTGGCAGCGTCAGCCAGGGGAGCATCAGTCAGGGGAGTACAGAGTACATCTCTGTACCG 703
Db 607 CTTTGGCAGCG-----TCAGTGAGGAGATGACAGATATATCTGTGACCG 651
Qy 704 CTTTACCCCGATAGCCTGTGATGGTGTTCATTCAGTCCAGCATATATAGTGGGTCT 763
Db 652 CTTCTACCCCAATGACTTGTGGGTGTGTTCAGTTTCAGCACATCATCTGGTGGCT 711
Qy 764 CATCTGCGCGGCATCGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
Db 712 TATCTCTGCTGTGATTTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 771
Qy 824 CTCGAAGGCGCACCAAGAGCGGCTCAAGAGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCT 883
Db 772 CTCGAAGGCGCACCAAGAGCGGCTCAAGAGCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCT 831
Qy 884 CTTTGGCTGCTGGCTGCCATATATGTTGGGATCAGATCGACTCTCTCTCTCTCTCTCTCTCT 943
Db 832 CTTGCGCTGTTGGCTGCCCTTACTACATTTGGGATCAGCATCGACTCTCTCTCTCTCTCTCT 891
Qy 944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTGCACAAGTGGATCTCTCTCTCTCTCT 1003
Db 892 AATCATCAAGCAAGGATGTGACTTTGAGAACACATGTGCACAAGTGGATTTCCATCACCGA 951
Qy 1004 GGCCTCGCTTCTTCCACTGTGCTGCTGAAACCCATCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
Db 952 GGCCTAGCTTCTTCCACTGTGCTGAAACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
Qy 1064 GTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCTCAAGAT 1123
Db 1012 ATTTAAACCTCTGCCAGCAGCAGCACTCACTCTGAGCAGAGGCTCCAGCTCAAGAT 1071
Qy 1124 CTTTCCAAAGAAAGCGGGTGGACACTCTTCCTGCTCCAGGAGTCCAGATCTCTCCAG 1183
Db 1072 CTTCTCCAAAGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG 1131
Qy 1184 TTTTCACTCCAGCTAAACCTTATGCAAGACTTATATAATATATATATATATATATATAA 1243
Db 1132 TTTTCACTCCAGCTAAACAGATGTAAAGACTTTT-----TTTATAGATAAATA 1183
Qy 1244 GAACCTTTTATGTTACACTTTCCAGATATAAGAGACTGACAGCTGTTGTACAGTTTT 1303
Db 1184 ACTTTTAAAGTTTACACTTTTTCAGATATAAAGACTGACCAATATTTGTACAGTTTT 1243
Qy 1304 TTTTCTTTTAAATGACTGTTGGGAGTTTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1363
Db 1244 TATTGCTTGTGGATTTTGT-----CTTGTGTTTCTTTAGTTTGTGAAGTTTAAAT 1297
Qy 1364 TAATTTATATAAATATTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1423
Db 1298 GACTTATTTATA-----TAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1353
Qy 1424 GTGGCCAGTTCTTAGTAGCTGTTTATCTGTGTGAGGACTGTAGAACTGTAGAGGAAGA 1483
Db 1354 GTGGCCAGTTCTTAGTTGCTGTATGCTCTGTTGGTAGGACTGTAGAA-----AAGGG 1405
Qy 1484 AACTGAACATTCAGAAATGTGGTAAATTTGAATTAAGCTAGCCGTGATCTCTCAGCTGTT 1543
Db 1406 AACTGAACATTCAGAGCGTGTAGTGAATCAGTTAAAGCTAGAAATGATCCCCAGCTGTT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473

```

; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van 't Veer, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-198-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-912

```

652	DB	CTTCTACCCCAATGACTTGTGGGGTGTGTGTCAGTTTCAGACACATCATGGTTGGCCT	711
764	QY	CATCTCGCCGGGCATCGTCATCCTCTCTGTGTTACTGCAATCATCATCTCTTAAGCTGTGCACA	823
712	DB	TATCTCGCTGCTGTATTGTTCATCCTGTGCTGTATTGCAATTAATCATCTCCAAAGCTGTGCACA	771
824	QY	CTCCAAAGGCCACCAAGAGCGCAAGCCCTCAAGACGACAGTCATCTTCATCCTCCTAGCTTT	883
772	DB	CTCCAAAGGCCACCAAGAGCGCAAGCCCTCAAGACCAAGTCATCTTCATCCTCCTGCTTT	831
884	QY	CTTTGCTCTGCTGGCTGCCATATTATGTGGGGATCAGCATCGACTCTTTTCATCTTTTGGG	943
832	DB	CTTTCGCTCTGTGGCTGCCCTTACTACATTGGGATCAGCATCGACTCTTTTCATCTCTCTGGA	891
944	QY	AGTCATCAAGCAAGGATGTGACTTCGAGAGCATGTGCACAAAGTGGATCTCCATCACAGA	1003
892	DB	AATCATCAAGCAAGGCTGTGAGTTTGAGAACACTGTGCACAAGTGGATTTCCATCACCGA	951
1004	QY	GGCCCTCGCCTCTTCCACTGTGTGCTCGAAACCCCATTCCTATGAGCTTCTCTGGGGCCAA	1063
952	DB	GGCCCTTAGCTTTCTTCCACTGTGTGTGTAACCCCATCTCTATGTTTCTCTTGGAGCCAA	1011
1064	QY	GTTCAAAGCTCTGCCACGATGCACTCAACTCCATGAGCAGAGGCTCCAGGCTCAAGAT	1123
1012	DB	ATTTAACAACCTCTGCCACGACGCACTCACTCTGTGACGAGAGGTCAGGCTCAAGAT	1071
1124	QY	CTTTTCCAAAGGAAGCGGGGTGGACACTCTTTCCTCTCCACGAGTCAAGATCTCTCAG	1183
1072	DB	CCTCTCCAAAGGAAGCGAGGTGGACATTCATCTGTTTCCAATGAGTCTGTGAGTCTCAAG	1131
1184	QY	TTTTCACTCCAGCTAACCCCTTATGCAAAAGACTTATAATAATATATATATATATGATAAA	1243
1132	DB	TTTTCACTCCAGCTAACACAGATGTAAAGACTTTTTT-----TTTATACGATAAATA	1183
1244	QY	GAACTTTTTTTATGTTTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTGTACAGTTT	1303
1184	DB	ACTTTTTTTTAAAGTACACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTT	1243
1304	QY	TTTTTTTTTTTTAAATGACTGTGTGGAGTTTATGTTCTCTAGTTTTCGTAGGTTTGACT	1363
1244	DB	TATGCTCTGTGGAATTTTGT-----CTTGTGTTTCTTTAGTTTTTGTGAAGTTTAAT	1297
1364	QY	TAATTTATATAATATTGTTTTTGTGTTTTCATGTGAATGAGCGCTGTAGCAGGACCT	1423
1298	DB	GACTTATTTATA-----TAAATTTTTTTTGTTCATATTGATGTGTCTAGCAGGACCT	1353
1424	QY	GTGGCCAAAGTCTTAGTAGTACTTTTATCTGTGTGAGGACTGTAGAAGCTGTAGAGGAAGA	1483
1354	DB	GTGGCCAAAGTCTTAGTGTCTGTATGTCTCGTGTGAGGACTGTAGAA-----AAGGG	1405
1484	QY	AACTGACATCCCAAGTGTGTGGTAAATTGGAATAAAGCTAGCCGTGATCCTCAGCTGTT	1543
1406	DB	AACTGACATCCCAAGGCGGTGTGTGAATCACGTAAGCTTGAATAATGATCCCAAGCTGTT	1465
1544	QY	GCTGCATA	1551
1466	DB	TATGCATA	1473

RESULT 6
US-10-151-274-14
; Sequence 14, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274

; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-151-274-14

Query Match 50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

QY 44 GCAGGTGACAGTACAGTACCTCTGAGGCGTTTGGTGTCTCGGTAAACACACGCGCTG 103
DB 12 GCGGACGACGATAGCAAGTACGCGGCGCTGAGTGTCTCCAGTAGCCACGCGATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACGATCAGTGTGAGTATATACATCTCTGATAACTACTC 163
DB 72 GAGAACGACGCGTTACCATGGA-----GGGATCAGTATATACATCTCAGATMACTAC 126
QY 164 TGAAGAGTGGGTCTCGAGACTATGACTCCAAAGGAACCTGCTTCGGGATGAAAA 223
DB 127 CGAGGAATGGCTCAGGGACTATGACTCCATGAGGAACCTGTTTCGTGAAGAAAA 186
QY 224 GGTCCATTGAAATAGATCTCTCGCCACGATCTACTTCACTATCTCTTGTGACTGGCAT 283
DB 187 TGCTAAATTTCAATAAAATCTCTCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 246
QY 284 AGTCGCAATGATTTGTTGATCTCTGCTGATGGGTTACGAGAAGAGCTAAGAGCATGAC 343
DB 247 TGTGGGCAATGATTTGTTGATCTCTGCTGATGGGTTACGAGAAGAGCTAAGAGCATGAC 306
QY 344 GGACAAGTACCGGTGACCTGTGACGTGAGTGTGACCTCTCTTTGTGATCACTCCGCTT 403
DB 307 GGACAAGTACCGGTGACCTGTGACGTGAGTGTGACCTCTCTTTGTGATCACTCCGCTT 366
QY 404 CTGGCAGTTGATCGATCGATGCTGATCTGTTGGGAAATTTTGTGTAAGCTGTCCA 463
DB 367 CTGGCAGTTGATCGATCGATGCTGATCTGTTGGGAAATTTTGTGTAAGCTGTCCA 426
QY 464 TATCATCTACACTGTCAACCTCTACAGCAGGTTTCTCATCTGCGCTTTCATCAGCCTGGA 523
DB 427 TGTCACTACAGTCAACCTCTACAGCAGTGTCTCATCTGCGCTTTCATCAGTCTGGA 486
QY 524 CCGGTACCTCGCATGTGTCAGGCAACCAAGTCAAGGCAAGGAACTGCTGGCTGA 583
DB 487 CCGGTACCTCGCATGTGTCAGGCAACCAAGTCAAGGCAAGGAACTGCTGGCTGA 546
QY 584 AAAGCAGTCTATGTGGCGTCTGATCCAGCCCTCTCTCTGACTATACCTGACTTCAT 643
DB 547 AAAGTGTCTATGTGGCGTCTGATCCCTGCTGCTCTCTGACTATTCGAGCTTCAT 606
QY 644 CTTTCCGACGTCACCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCG 703
DB 607 CTTTCCCAACG-----TCAGTCAGGCGATGACAGATATATCTGTGACCG 651
QY 704 CTTTACCCGATAGCTGTGATGCTGTTTCAATTCAGCATATAATGCTGGCTCT 763
DB 652 CTTTACCCGATAGCTGTGATGCTGTTTCAATTCAGCATATAATGCTGGCTCT 711
QY 764 CATCTGCGCGCATCTGATCTCTCTGTTTACTGCAATCATCTCTAAGCTGTGACA 823
DB 712 TATCTGCTGGTATTTGTCATCTCTGCTGCTATTTGCAATATCATCTCCAAGCTGTGACA 771
QY 824 CTTCAAGGCGCACAGAGCGAAGGCGCTCAGACGACGATCTCTCATCTCTAGCTTT 883
DB 772 CTTCAAGGCGCACAGAGCGAAGGCGCTCAGACGACGATCTCTCATCTCTAGCTTT 831
QY 884 CTTTCTGCTGCTGCTGCTATTTATGTTGGGATCAGCATCTCTCTCTCTTTTGGG 943

DB 832 CTTGCGCTGTGGCTGCTTACTACATTTGGATCAGCATCGACTCTCTCTCTCTGGA 891
QY 944 AGTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTCACAAGTGGATCTCCATCAGA 1003
DB 892 AATCATCAAGCAAGGTTGAGTTTGAACACATGTGCACAAGTGGATTTCCATCACCAG 951
QY 1004 GGCCTCGCCTTCTTCCACTGTTCCTGAAACCCCATCTCTATGCTCTCTCTCTCTGAGGCA 1063
DB 952 GGCCTAGCTTCTTCCACTGTTCCTGAAACCCCATCTCTATGCTCTCTCTCTCTGAGGCA 1011
QY 1064 GTTCAAAAGCTCTCCGACGATGCACTCACTCATGAGCAGGCTCCAGCCTCAAGAT 1123
DB 1012 ATTTAAACCTCTGCCCAGCAGCACTCACTCTGTGAGCAGAGGCTCCAGCCTCAAGAT 1071
QY 1124 CTTTCCAAAGAAAGCGGGGTGACACTCTTCCGTCTCCACGAGAGTCAGAATCCTCCAG 1183
DB 1072 CTTCTCCAAAGAAAGCGGGGTGACACTCTCTGTTTCCACTGAGTCTGAGTCTTCAAG 1131
QY 1184 TTTTCACTCCAGCTAAACCTTTATGCAAGACTTATATATATATATATATATATATATATA 1243
DB 1132 TTTTCACTCCAGCTAAACAGATGTAAAGACTTTT-----TTTATACGATAATA 1183
QY 1244 GAACTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCACTCTCTGACAGTTT 1303
DB 1184 ACTTTTATGTTTACACATTTTTCAGATATAAAGACTGACCAATATTTGACAGTTT 1243
QY 1304 TTTTATTTTAAATGACTGTGGAGTTTATGTTCTCTAGTTTCTGAGGTTTGA 1363
DB 1244 TATTCCTTGTGGATTTTGT-----CTTGTGTTTCTTTAGTTTCTGAGTTTAAAT 1297
QY 1364 TAAATTTATATAATATTTGTTTGTGTTTCTGTAATGAGGCTCTAGCAGGACT 1423
DB 1298 GACTTATTTATA-----TAAATTTTTTTTGTTCATATTTGATGTTGTTAGCAGGACT 1353
QY 1424 GTGGCAAGTCTTCTAGTAGCTTTTATCTGTGTGAGGACTGTAGAACTGTAGAGAGA 1483
DB 1354 GTGGCAAGTCTTCTAGTTGCTGATGTTCTGTTGTTAGGACTGTAGAA-----AAGGG 1405
QY 1484 AACTGAACATTCAGAAATGTGTGTTAAATTTGAATTAAGCTAGCCGTGATCTCTCAGCTGT 1543
DB 1406 AACTGAACATTCAGAGGCTGTAGTGAATCAGTAAAGCTAAGAAATGATCCCCAGCTGT 1465
QY 1544 GCTGCATA 1551
DB 1466 TATGCATA 1473

RESULT 7

US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003467

; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-912

Query Match 50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

```
QY 44 CGAGTGAGGTAGGAGTGCCTCTGAGGCGTTTGGTCTCCGGTAACCAACGCGCTG 103
DB 12 CGGCGAGCAGGTAGCAAAAGTACGCCGAGGCGCTGAGTGCTCCAGTAGCCCGCATCTG 71
QY 104 TAGAGCGAGTGTGCGCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTCTC 163
DB 72 GAGAACAGCGGTTACCATGA-----GGGATCAGTATATACACTTCAGATAACTACAC 126
QY 164 TGAAGAGTGGGGTCTGAGACTATGACTCCAAAGGAACCCCTGCTCCGGATGAAAA 223
DB 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGAACCCCTGTTCCGTGAAGAAAA 186
QY 224 GTCCTATTTCATAGATGCTTCTCCGCCACCATCTACTTTCATCATCTTCTTGACTGGCAT 283
DB 187 TGCTAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTTTAACTGGCAT 246
QY 284 AGTCGGCAATGATTTGGTATCTCTGTCATGCGTTTACAGGAAGCTTAAGGACATGAC 343
DB 247 TGTGGGCAATGATTTGGTATCTCTGTCATGCGTTTACCGAAGAACTGGAAGCATGAC 306
QY 344 GAGCAAGTACCGCTGCACTGTGACGTGCTGACCTCTCTTTGTTCATCACACTCCCTT 403
DB 307 GGACAAGTACAGCTGCACTGTGACGTGCGGACCTCTCTTTGTTCATCACGCTTCCCTT 366
QY 404 CTGGGCGAGTTGATGCGATGCTGCTGACTGCTGTTTGGGAAATTTTGTGTAAGGCTGCCA 463
DB 367 CTGGGCGAGTTGATGCGGTGGCAAACTGGTACTTTGGGAACTTCTTATGAAAGGAGTCCA 426
QY 464 TATCATCTACATGCTCAACTCTACAGCAGCTTCTCATCTCGGCTTTCATCAGCTGGA 523
DB 427 TGTATCTACACAGTCAAACTCTACAGCAGTGTCTCATCTCGGCTTTCATCAGTCTGGA 486
QY 524 CGGTAACCTCGCATTTGTCAGGCCACCAACAGTCAAGGCCAAGGAAACTGCTGGCTGA 583
DB 487 CGGTAACCTCGCATTTGTCAGGCCACCAACAGTCAAGGCCAAGGAAAGCTTGTGGCTGA 546
QY 584 AAGGCGAGTCTATGTTGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTAT 643
DB 547 AAGGTTGGTCTATGTTGGGCTGTGATCCCTCTCTGCTGACTATTTCCGACTTCAT 606
QY 644 CTTTTCGCGAGTCAAGCCAGGGGACATCAGTCAGGGGATGACAGTACATCTGTGACCG 703
DB 607 CTTTTCGCAACG-----TCAGTGGGCGAGATGACAGATATATCTGTGACCG 651
QY 704 CTTTTCGCGAGTCAAGCCAGGGGACATCAGTCAGGGGATGACAGTACATCTGTGACCG 763
DB 652 CTTTTCGCGAGTCAAGCCAGGGGACATCAGTCAGGGGATGACAGTACATCTGTGACCG 711
QY 764 CATCTCTCGCGGATGCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823
DB 712 TATCTCTCGCGGATGCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771
QY 824 CTCAAGGGGACCAAGAGCGCAAGGCCCTCAAGACGACAGTCACTCTCATCTCTAGCTTT 883
DB 772 CTCAAGGGGACCAAGAGCGCAAGGCCCTCAAGACGACAGTCACTCTCATCTCTAGCTTT 831
QY 884 CTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
DB 832 CTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
QY 944 AGTCATCAAGCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
DB 892 AATCATCAAGCAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
QY 1004 GGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
```

```
DB 952 GGCCCTAGCTTTCTTCCACTGTGTCTGAAACCCCATCTCTATGCTTTCTCTTGAGCCAA 1011
QY 1064 GTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAGGCTCCAGCCTCAAGAT 1123
DB 1012 ATTAAAAACCTCTGCCAGCAGCGCACTCACTCTCTGAGCAGAGGGTCCAGCCTCAGAT 1071
QY 1124 CTTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTCTCTCCAGGAGTCAAGATCTCTCCAG 1183
DB 1072 CTTTCCAAAGGAAAGCGAGGTGGACATTCATCTGTTTCCACTGAGTCTGAGTCTTCAAG 1131
QY 1184 TTTTCACTCCAGCTAACCCCTTATGCAAGACTTATATATATATATATATATATATATATA 1243
DB 1132 TTTTCACTCCAGCTAACCAAGACTTTT-----TTTATACGATAAATA 1183
QY 1244 GAACCTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCAAGTCTTGTACAGTTTT 1303
DB 1184 ACTTTTAAAGTTACACATTTTTCAGATATAAAGACTGACCAATATTTGACAGTTTT 1243
QY 1304 TTTTAAAGTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1363
DB 1244 TATTGCTTCTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1297
QY 1364 TAATTTATATAAATATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1423
DB 1298 GACTTATTTATA-----TAAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1353
QY 1424 GTGGCCAAAGTCTTACTGACTGCTGTTTATCTGTGTGAGGACTGTAGACTGTAGAGGAAGA 1483
DB 1354 GTGGCCAAAGTCTTACTGACTGCTGTTTATCTGTGTGAGGACTGTAGAA-----AAGGG 1405
QY 1484 AACTGAACATTCAGAAATGTTGGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1543
DB 1406 AACTGAACATTCAGAAATGTTGGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1465
QY 1544 GCTGCATA 1551
DB 1466 TATGCATA 1473
```

RESULT 8
US-10-170-385-332
; Sequence 332, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 53268200100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-332

Query Match 50.8%; Score 952.8; DB 13; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

1124	QY	CTTTCCAAAGGAAGGGGTGGACACTCTTCGGTCTCCACGGAGTCAGAAATCCTCCAG	1183
1072	Db	CCTCTCCAAAGGAAGCGAGGTGGACATTCATCTGTTTCCACTGAGCTGAGCTTCAAG	1131
1184	QY	TTTTCACTCCAGCTAAACCCCTTATGCAAGACTTATATAATATATATATATATGATATAA	1243
1132	Db	TTTTCACTCCAGCTAACACAGATGTAAGACTTTTT-----TTTATACGATAAATA	1183
1244	QY	GAACTTTTTTATGTTTACACATTTTCCAGATATAAGAGACTGACCAAGCTCTGTACAGTTTT	1303
1184	Db	ACTTTTTTTTAAAGTTTACACATTTTTCAGATATAAAGACTGACCAATATTTGTACAGTTTT	1243
1304	QY	TTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCTCTAGTTTCTGTAGGTTTGACT	1363
1244	Db	TATTGCTGTGTGGATTTTGT-----CTTGTTTTCTTTAGTTTTGTGAATTTAAT	1297
1364	QY	TAAATTATATAATAATCTTTTTTTTTGTTTTTCATGCGAATGAGCGTCTAGCAGGACCT	1423
1298	Db	GACTTATTATATA---TAAATTTTTTTTTGTTTTTCAATTGAATGTGTGCTTAGGCAGGACCT	1353
1424	QY	GTGGCCAAAGTTCTTAGTAGCTGTTTATCTGTGTGTAGGACTGTAGAACTGHPAGAGGAAGA	1483
1354	Db	GTGGGCCAAGTCTTAGTTAGTTGCTGTATGTTCTGCTGGTAGGACTGTAGAA-----AAGGG	1405
1484	QY	AACTCAACATTCACGAATGTGTGGTAAATTGAATAAAGCTAGCCCGTGATCCTCAGCTGTT	1543
1406	Db	AACTGAAACATTCACAGAGCGGTAGTGAATCAGTAAAGCTAGAAATGATCCCCAGCTGTT	1465
1544	QY	GCTGCATA	1551
1466	Db	TATGCATA	1473

```

Db      1244  TAATGCTCTTTGGATTTTGT-----CTGTGTTTCTTTAGTCTTTTGTGAAGTTTAATT 1297
QY      1364  TAATTTATATAAATATTGTTTTTTTGTGTTTTCATGTGAATGAGCGTCTAGGCAGGAACCT 1423
Db      1298  GACTTATTTTATA-----TAAATTTTTTTTGTTCATATTGATGTGTCTTAGGCAGGACCT 1353
QY      1424  GTGGCCAAAGTCTTCTAGTACTGTTTATCTGTGTGTAGGACTCTAGAACTGTAGAGGAAGA 1483
Db      1354  GTGGCCAAAGTCTTCTAGTCTGTATGTCGTGTGGTAGGACTGTAGAA-----AAGGG 1405
QY      1484  AACTGAACATTCAGAAATGTGTGGTAAATTAAGCTAGCCGTGATCCTCAGCTGTT 1543
Db      1406  AACTGAACATTCAGAGCGGTAGTAGAATCACGTAAGACTAGAAATGATCCCCAGCTGTT 1465
QY      1544  GCTGCATA 1551
Db      1466  TATGCATA 1473

RESULT 10
US-10-021-660-58
; Sequence 58, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-58

Query Match          50.8%; Score 952.8; DB 15; Length 1679;
Best Local Similarity 80.2%; Pred. No. 5e-209; Indels 46; Gaps 6;
Matches 1210; Conservative 0; Mismatches 252;

QY      44  GCAGGTGCAGGTAGCAGTGCACCTCTGTAGGCGGTTTGGTGTCTCGGTAACCAACCAACGCGCTG 103
Db      12  GCGGCACGAGTAGCAAAAGTGACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
QY      104  TAGGCGAGTGTGCCATGGAACCGATCAGTGTAGTATATACACTTCTGTAACTACTC 163
Db      72  GAGAACACGCGGTTTACCATGGA-----GGGGATCAGTATATACACTTTCAGTAACTACAC 126
QY      164  TGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAGAAACCCCTGCTTCGGGATGAAAA 223
Db      127  CGAGGAATGGCTCAGGAGCACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAAA 186
QY      224  CGTCCATTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCTTCTTGACTGGCAT 283
Db      187  TGC'TAATTTCAATAAATATCTTCTGCGCCACCATCTACTTCATCATCTTCTTAACTGGCAT 246
QY      284  AGTCGGCAATGGATTGTGTGATCTGTGTCATGGGTATACCAAGAAAGCTAAGGAGCATGAC 343
Db      247  TGTGGGCAATGGATTGTGTATCTTGTGTCATCTTGTCTATGGGTACCAGAAGAACTGAGAAGCATGAC 306
QY      344  GGACAAGTACCGGTGCACCTGTCAGTGGTGCACCTCTCTTTGTGATCACTCCCTT 403
Db      307  GGACAAGTACAGGTGCACCTGTCAGTGGCGCACCTCTCTTTGTGATCACTCCCTT 366

```

Qy 404 CTGGGAGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGCTGTCCA 463
Db 367 CTGGGAGTTGATGCCGCTGGCAACTGGTACTTTGGGAATTTCTTATGCAAGGAGTCCA 426
Qy 464 TATCATCTACAGCTGTCAACCTCTACAGCAGGCTTCATCTCGGCTTCATCAGCCTGGA 523
Db 427 TGTCTATCTACAGCTGAACCTCTACAGCAGTTCCTCATCTGGCTTCATCAGTCTGGA 486
Qy 524 CCGGTACTCTGCCATTTCCACGCCCAACACAGCTCAAGGCCCAAGAACTGTGCTGGA 583
Db 487 CCGGTACTCTGCCATCTGTCCACGCCCAACACAGCTCAAGGCCCAAGAACTGTGCTGGA 546
Qy 584 AAAGGAGTCTATGTGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT 643
Db 547 AAAGGAGTCTATGTGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT 606
Qy 644 CTTTGGCGAGCTCAGCCAGGCGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCG 703
Db 507 CTTTGGCGACG-----TCAGTGAAGGAGATGACAGATATATCTGTGACCG 651
Qy 704 CTTTACCCGAGTCTGTGGATGGTGGTGTTCATATTCAGCATATATGTGGGTCT 763
Db 652 CTTTACCCGAGTCTGTGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT 711
Qy 764 CATCTGCCGGCATCTGCTCTCTCTGTTACTGATCATCATCTTAAGCTGTCA 823
Db 712 TATCTGCTGCTGTATGCT 771
Qy 824 CTCGAAGGCCCAACAGAGCCCAAGGCCCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCT 883
Db 772 CTCGAAGGCCCAACAGAGCCCAAGGCCCTCAAGCAGCAGTCTCTCTCTCTCTCTCTCT 831
Qy 884 CTTTGGCTGCTGGTGCATATATATGTTGGGATCAGATCAGCTCTCTCTCTCTCTCTCT 943
Db 832 CTTTGGCTGCTGGTGCATATATGTTGGGATCAGATCAGCTCTCTCTCTCTCTCTCTCT 891
Qy 944 AGTATCAAGCAAGATGTGCTCGAGCATGTTGCAAGTGGATCTCCATCACAGA 1003
Db 892 AATCATCAAGCAAGGTTGAGTGTGAGAACACTGTGCAAGTGGATTTCCATCACCGA 951
Qy 1004 GGCCCTGCTCTTCTTCCACTGTTGTTGCAAGCCCATCTCTCTCTCTCTCTCTCTCTCT 1063
Db 952 GGCCCTGCTCTTCTTCCACTGTTGTTGCAAGCCCATCTCTCTCTCTCTCTCTCTCTCT 1011
Qy 1064 GTTCAAGAGCTCTGCCAGCATGCACTCAATCTCATGAGCAGAGGCTCCAGCCTCAAGAT 1123
Db 1012 ATTTAAAGCTCTGCCAGCAGCAGCTCACTCTCTGTGAGCAGAGGCTCCAGCCTCAAGAT 1071
Qy 1124 CTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGCTCTCCAGGAGTCAAGATCTCTCAG 1183
Db 1072 CTTTCCAAAGGAAAGCGGGGTGGACACTCTTCCGCTCTCCAGGAGTCAAGATCTCTCAG 1131
Qy 1184 TTTTCACTCCAGCTAAACCTTATGCAAGACTTATATATATATATATATATATATATATA 1243
Db 1132 TTTTCACTCCAGCTAAACAGATGTAAGAGACTTTTT-----TTTATACGATTAATA 1193
Qy 1244 GAACTTTTTATGTTACACATTTCCAGATATAGAGACTGACAGTCTTGTACAGTTTT 1303
Db 1184 ACTTTTTTTTAAAGTTTACACATTTTTCAGATATAAAGACTGACCAATATGTACAGTTTT 1243
Qy 1304 TTTTCTTTTAAATGACCTGTGGGAGTTATGTTCTCTAGTTTCTGAGGTTGACT 1363
Db 1244 TATGCTGTTGGATTTTGT-----CTTGTGTTCTTTAGTTTCTGAGGTTTAAAT 1297
Qy 1364 TAATTATATAATATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1423
Db 1298 GACTTATTTATA-----TAAATTTTTTTTGTGTTTCAATTTGATGTTGTTGTTGTTGTT 1353
Qy 1424 GTGGCAAGTTCTTATAGTCTTTTATCTGTGTGTAGGACTGTAGAACTGTAGAGGAAGA 1483
Db 1354 GTGGCAAGTTCTTATAGTCTTTTATCTGTGTGTAGGACTGTAGAA-----AAGGG 1405

Qy 1484 AACTGAACATTCAGAAATGTGTGTGAATTAATTAATAAAGCTAGCCGTGATCTCAGCTGTT 1543
Db 1406 AACTGAACATTCAGAGCGTGTAGTGAATACGTAAAGCTAGAAATGATCCCAAGCTGTT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473
RESULT 11
US-10-341-434-177
; Sequence 177, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 90 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 177
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(1144)
; OTHER INFORMATION:
US-10-341-434-177

Query Match 50.8%; Score 952.8; DB 16; Length 1679;
Best Local Similarity 80.2%; Pred. No. Se-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;
Qy 44 GCAGGTGACAGTACAGTACACCTCTGAGGCGCTTTGGTGTCTCCGGTAACCAACCAAGGCTG 103
Db 12 CGGCGACAGGTAGCAAGTACGCGGAGGCTCAGTGTCCAGTAGGCAACCGCATCTG 71
Qy 104 TAGAGCGAGTGTGCCATGGAAACCGATCAGTGTGAGTATATACACTTCTGTAACTACTC 163
Db 72 GAGAACCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCCAGTAACATAC 126
Qy 164 TGAAGAAGTGGGCTCTGGAGACTATGACTCCAAACAGGAACCTCTCTCCGGGATGAAAA 223
Db 127 CGAGGAATGGGCTCAGGGACTATGACTCCATGAGGAACCTCTCTCCGTGAAGAAAA 186
Qy 224 CGTCCATTTCAATAGGATCTTCTGCCCCACCATCTACTTCATCATCTTCTGACTGGCAT 283
Db 187 TGCTAATTTCAATAAATCTTCTGCCCCACCATCTACTTCATCATCTTCTTAACTGGCAT 246
Qy 284 AGTCGGCAATGGATTTGGTGTCTGCTGATCTGCTGATGGTTACCAAGAGAGCTAAGGAGCATGAC 343
Db 247 TGTGGCAATGGATTTGGTGTCTGCTGATCTGCTGATGGTTACCAAGAGAGCTAAGGAGCATGAC 306
Qy 344 GGAACAAGTACCGGCTGCACCTGTCTGAGTGGCTGACCTCTCTCTTTGTCTCATCACACTCCCTTT 403
Db 307 GGAACAAGTACAGCTGCACCTGTCTGAGTGGCTGACCTCTCTCTTTGTCTCATCACACTCCCTTT 366
Qy 404 CTGGCAGTTGATGCGCATGGCTGACTGGTACTTTGGGAAATTTTGTGAAGCTGTCCA 463
Db 367 CTGGCAGTTGATGCGCTGGCAAACTGGTACTTTGGGAACTTCCATATGCAAGGAGCTCCA 426
Qy 464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCTCCCTGCTTCATCAGCCTGGA 523
Db 427 TGTCTATCTACAGTCAACCTCTACAGCAGTGTCTCTCATCTCCCTGCTTCATCAGTCTGGA 486
Qy 524 CCGGTACTCTGCCATTTGTCCACGCCCAACAGCTCAAGGCCCAAGAACTGTGCTGCTGA 583

Db 487 CGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGGCCAAGAAAGCTGTTGGCTGA 546
Qy 584 AAAGCAGCTATGTGGGGCTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCAT 643
Db 547 AAAGTGGTCTATGTTGGGGCTCTGGATCCCTGCTCTCTGCTGACTATTCGCGACTTCAT 606
Qy 644 CTTTGGCGAGCTCAGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCG 703
Db 607 CTTTGGCAACG-----TCAGTGAGGCGAGATGACAGATATCTGTGACCG 651
Qy 704 CTTTACCOCGATGAGCTCTGGATGGTGTGTTTCAATTCCAGCATATAAAGTGGGTCT 763
Db 652 CTTTACCOCGATGAGCTCTGGTGGTGTGTTTCCAGTTTCAGCATCATGTTGGCT 711
Qy 764 CATCTGCGCGGATCGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 823
Db 712 TATCTCTGCTGATGTTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 771
Qy 824 CTCGAAGGGCCACCAAGGCGCAAGCCCTCTCAAGCAGCAGTCATCTCTCTCTCTCTCTCT 883
Db 772 CTCGAAGGGCCACCAAGGCGCAAGCCCTCTCAAGCAGCAGTCATCTCTCTCTCTCTCTCT 831
Qy 884 CTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943
Db 832 CTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
Qy 944 AGTCATCAAGCAAGGATGAGCTCTCAGAGCAGTGTGCAAGAGTGTCTCTCTCTCTCTCTCA 1003
Db 892 ATTCATCAAGCAAGGATGAGCTCTCAGAGCAGTGTGCAAGAGTGTCTCTCTCTCTCTCTCA 951
Qy 1004 GGCCCTGCGCTCTCTCACTGTTGCTGCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
Db 952 GGCCCTGAGCTCTCTCTCACTGTTGCTGCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
Qy 1064 GTTCRAAGCTCTGCCAGATGATCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123
Db 1012 ATTAAACCTCTGCCAGCAGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1071
Qy 1124 CTTTCCAAAGGAAGCGGGGTGGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
Db 1072 CTTCTCAAGGAAGCGAGGTGGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1131
Qy 1184 TTTTCACTCCAGCTTAACCTTTAGCAAGCTTATATATATATATATATATATATATATATATA 1243
Db 1132 TTTTCACTCCAGCTTAACAGATGTAAAGACTTTT-----TTTATACGATTAATA 1183
Qy 1244 GAACCTTTTATGTTACACATTTTCCAGATATAAGAGACTGACCACTCTTGTACAGTTT 1303
Db 1184 ACTTTTATTTTATGTTACACATTTTCCAGATATAAGAGACTGACCACTATTTGTACAGTTT 1243
Qy 1304 TTTTATTTTATTTTAAATGCTGTTGGAGTTTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1363
Db 1244 TATTGCTTGTGGATTTTGT-----CTTGTGTTTCTTTAGTTTGTGAAAGTTTAAAT 1297
Qy 1364 TAATTTATATAATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1423
Db 1298 GACTTATATA-----TAATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1353
Qy 1424 GTGCCAAGTTCTTATGAGTGTGTTTATCTGTGTGTAGGACTGTAGAACTGTAGAGAGA 1483
Db 1354 GTGCCAAGTTCTTATGTTGCTGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1405
Qy 1484 AACTGAACATCCAGATGTGTGTAATTTGAATAAGCTAGCGGTGATCTCTCTCTCTCTCTCT 1543
Db 1406 AACTGAACATCCAGAGCGTGTAGTGAATCACGTAAGCTAGAAATGATCCCCAGCTGTT 1465
Qy 1544 GCTGCATA 1551
Db 1466 TATGCATA 1473

RESULT 12

US-09-971-392-20

; Sequence 20, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 284616.2
; NAME/KEY: unsure
; LOCATION: 1706
; OTHER INFORMATION: a, t, c, g, or other
; US-09-971-392-20

Query Match 50.8%; Score 952.8; DB 10; Length 1711;
Best Local Similarity 80.2%; Pred. No. 5e-209;
Matches 1210; Conservative 0; Mismatches 252; Indels 46; Gaps 6;

Qy 44 GCAGTGCAGGTAGCAGTGCACCTCTGAGGGCTTGTGCTCCGGTAACCAACGCGCTG 103
Db 47 GCGGCACGAGGTAGCAAGTAGCGCGAGGGCTGAGTGTCCAGTAGCCACCGCATCTG 106
Qy 104 TAGAGCAGATGTTCCCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATACTC 163
Db 107 GAGAACGACGCTTACCATGGA-----GGGATCAGTATATACACTTCAGATACTAC 161
Qy 164 TGAAGAGTGGGGTCTGGAGACTATGACTCAACAGAAACCTGCTCCGGATGAAAA 223
Db 162 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAAACCTGTTTCGTGAGAAAA 221
Qy 224 CGTCCATTTCAATAGGATCTTCCTGCCACCATCTACTTTCATCATCTTCTTACTGGCAT 283
Db 222 TGCTAATTTCAATAAATCTTCTGCCACCATCTACTCCATCATCTTCTTAACTGGCAT 281
Qy 284 AGTCGGCAATGATGTGTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 343
Db 282 TGTGGCAATGATGTGTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 341
Qy 344 GGACAGTACCGGCTGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 403
Db 342 GGACAGTACCGGCTGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 401
Qy 404 CTGGGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
Db 402 CTGGGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
Qy 464 TATCATCTACACTGTCAACCTCTACAGAGGTTCTCATCTCTGCTGCTGCTGCTGCTGCTG 523
Db 462 TGTCTCTACACAGTCAACCTCTACAGAGGTTCTCATCTCTGCTGCTGCTGCTGCTGCTG 521
Qy 524 CGGTAACCTCGCCATGTCACGCGCCACCAACAGTCAAGGCAAGGAACTGCTGGCTGA 583
Db 522 CGCTACCTCGCCATGTCACGCGCCACCAACAGTCAAGGCAAGGAACTGCTGGCTGA 581
Qy 584 AAAGCAGTCTATGTGGCGTCTGATCTCCAGCCCTCTCTCTGACTATACCTGACTTCAT 643
Db 582 AAAGTGTCTATGTGGCGTCTGATCTCCAGCCCTCTCTCTGACTATATCCCGACTTCAT 641
Qy 644 CTTTCCGACGTCAGCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTCTGACCG 703
Db 642 CTTTCCCAACG-----TCAGTGAGGCGAGATGACAGATATATCTGTGACCG 686

	D	b		778	AAGGCCCCCAAGACCACAGTCATCTCATCCCGGTCTTCTTGCCCTGTGGCTGCCTTAC	837
	Q	y		906	TATGTGGGATCAGCATCGACTCCTTTCTATCCTTTTGGGAGTCATCAAGCAAGGATGTGC	965
	D	b		838	TACATTGGGATCAGCATCGACTCCTTATCCTCTCGAAATCATCAAGCAAGGATGTGC	897
	Q	y		966	TTTCGAGAGCAATTGTGGAAGAATGGATCTCCATCACAGAGGCCCTCGCCTCTTCCACTGT	1025
	D	b		898	TTTGGAAACATGTGTGCACAAGTGGATTTCCATCACGAGGGCCTAGCTTTCTTCCACTGT	957
	Q	y		1026	TGCTGAACCCCATCCTCTATGCCCTTCCTCGGGCCAAGTTCAAAGCTCTGCCCAGCAT	1085
	D	b		958	TGCTGNACCCCATCCTCTATGCTTCTTGAGCCNAATTAAACCTCTGCCCAGCAC	1017
	Q	y		1086	GCACTCAATCCATGAGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAAGCGGGGT	1145
	D	b		1018	GCACTCACCTCTGTGAGCAGAGGGTCCAGCCTCAAGATCCTCTCCAAAGGAAAGCGAGGT	1077
	Q	y		1146	GGACACTCTTCGCTCCACGAGGTCAGAACTCCTCCAGTTTTTCATCTCCAGCTTAACCCCTA	1205
	D	b		1078	GGACATCTATCTGTTCCACTGAGCTGAGTCTTCAAGTTTTTCATCTCCAGCTAA-----	1131
	Q	y		1206	TGCAAGACTTATATAATATATATATATATATATGATAAGAACCTTTTTATGTTACACATT	1265
	D	b		1132	--CACAGATGTAAGACACTTTTTTTATACGATAAATTAACTTTTTTTAAAGTTACACATT	1189
	Q	y		1266	TTCCAGATATAGAGACTCACAGCTCTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGT	1325
	D	b		1190	TTTCAGATATAAAGACTGACCAATATGTGACAGTTTTTATTGCTT---GTTGGATTTT	1245
	Q	y		1326	GGAGATTATGTTCCTCTAGTTTTTGTGAGTTTGACTTTAAATTTATATAAATPATTTGTTT	1385
	D	b		1246	TGCTCTTGTTCTTTTAGTTTTTCTGGAAGTTTAAATGACTTANTATATA---AAAT	1302
	Q	y		1386	TTGTTTTGTTTCATGTGAATGAGCGCTAGCCAGGACCTGTGCCAAGTCTTAGTAGCTG	1445
	D	b		1303	TTTTTTTGTTCATPATTGATGTGTCTAGCGAGACCTGTGCCAAGTTCCTTAGTTGCTG	1362
	Q	y		1446	TTTATCTGTGTAGCACTGTAGAACTGTAGAGGAAGAACTGAACATTCAGAAATGTGT	1505
	D	b		1363	TATGCTCTCGTGTAGGACGTAGAA-----AAGGGAACGMACTTCCAGAGCGTGT	1414
	Q	y		1506	GGTAAATTGAATAAAGTAGCCGTGATCTTCAGCTGTTGCTGCATATCTCTTCATTCCG	1565
	D	b		1415	AGTGAATCAGTTAAGCTAGAAAATGATCCCCAGCTGTTTATGCATAGATAATCTCTCCA-	1473
	Q	y		1566	AGGAGCACCCCACCCACCCACCCACCCATCTTAAATGTTTGGTTATGCTGT	1625
	D	b		1474	-----	1473
	Q	y		1626	GTGATGTTTGTGTGTTTTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1685
	D	b		1474	-----TTCCCGTGGAAACGTTTTTCTGTTCTTAAGAAGCTGATTTTGTGTAGAAGATG	1526
	Q	y		1586	GCATTTAAACCAAGCCTGAATGGTGGTAGAAATGCTGGGGTTTTTTTTTGTGTGTGTGT	1745
	D	b		1527	GCATTTAATACCAAGCCCAAGTGGT-ATGAAATGCTGG-----	1566
	Q	y		1746	TTTTTTTTCAGTTTTCAAGAGTAGATTGACTTCAGTCCCTACAAATGTACAGTCTGTGATTA	1805
	D	b		1567	-TTTTTTCAGTTTTTCAGAGTGGGTGATTTTCAGCACCTAC-AGTGTACAGTCTGTGATTA	1624
	Q	y		1806	CATTGTTAATAAAGTCAATGATTAACCTTAAAAAARAAAAAAAAA	1851
	D	b		1625	AGTTGTTAATAAAGTACATGTTAACTTAAAAAARAAAAAAAAA	1670

Search completed: May 17, 2004, 22:14:00
Job time : 934.766 secs

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:50:31 ; Search time 5417.72 Seconds
(without alignments)
10345.919 Million cell updates/sec

Title: US-09-367-052-1
Perfect score: 1877
Sequence: 1 ccatcctaatacactcact.....aaaaaaaaaaaaaaaaaaaa 1877

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	747.8	39.8	906	13	BU523025
2	720.4	38.4	1091	13	BU523025
3	707	37.7	1201	13	BU523025
4	705.6	37.6	1201	13	BU523025

5	704.6	37.5	1201	13	BU523025
6	701	37.3	1201	13	BU523025
7	685	36.5	1201	13	BU523025
8	684.8	36.5	1051	12	BU523025
9	660.8	35.2	1201	13	BU523025
10	647.4	34.5	1201	13	BU523025
11	631.4	33.6	958	12	BU523025
12	626.6	33.4	1201	13	BU523025
13	621.6	33.1	793	9	BU523025
14	621.2	33.1	639	9	BU523025
15	621.2	33.1	952	13	BU523025
16	610.2	32.5	753	12	BU523025
17	602.2	32.1	872	14	BU523025
18	599.2	31.9	945	13	BU523025
19	594	31.6	910	12	BU523025
20	593.4	31.6	911	13	BU523025
21	590	31.4	955	12	BU523025
22	578.4	30.8	891	12	BU523025
23	576.8	30.7	950	13	BU523025
24	576.2	30.7	628	10	BU523025
25	576	30.7	904	13	BU523025
26	574.2	30.6	886	13	BU523025
27	572	30.5	600	13	BU523025
28	566.2	30.2	928	14	BU523025
29	560.6	29.9	581	14	BU523025
30	558	29.7	1034	12	BU523025
31	552.8	29.5	765	12	BU523025
32	551.6	29.4	582	12	BU523025
33	550.2	29.3	876	14	BU523025
34	545.6	28.3	749	14	BU523025
35	549	29.2	1099	13	BU523025
36	547.4	29.2	827	12	BU523025
37	539.4	28.7	1201	13	BU523025
38	532.4	28.4	1201	13	BU523025
39	524.8	28.0	528	12	BU523025
40	524.8	28.0	750	12	BU523025
41	520.4	27.7	875	12	BU523025
42	519.2	27.7	976	13	BU523025
43	511.4	27.2	585	10	BU523025
44	509.6	27.1	913	12	BU523025
45	509	27.1	886	10	BU523025

ALIGNMENTS

RESULT 1
BU523025
LOCUS
DEFINITION
AGENCY: 10154335 NCI CGAP Co24 Mus musculus CDNA clone
IMAGE: 6525663 5', mRNA sequence.
ACCESSION
BU523025
VERSION
BU523025.1 GI:22833475
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 906)
NIH-MGC http://mgi.nhlbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://imgc.llnl.gov
Plate: LLAM14130 row: c column: 07
High quality sequence stop: 668.

```
FEATURES
source
Location/Qualifiers
1..906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529663"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_C024"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 39.8%; Score 747.8; DB 13; Length 906;
Best Local Similarity 97.3%; Pred. No. 1.9e-76;
Matches 795; Conservative 0; Mismatches 12; Indels 10; Gaps 3;

QY 52 AGGTAGCAGTACCTCTGAGCGTTTGTGCTCGGTAAACACACACCGCTGTAGACGA 111
DB 1 ATGTACACAGTACCTCTGAGCGTTTGTGCTCGGTAAACACACACCGCTGTAGACGA 60
QY 112 GTTTGCCATGAACCGATCACTGTGAGTATATACATCTTGATTAATCTCTGAAGAAG 171
DB 61 GTTTGCCATGAACCGATC-----AGTATATACATCTTGATTAATCTCTGAAGAAG 114
QY 172 TGGGCTCTGGACATGATCTCCACACAGGACCTGTTCCGGATGAAACGTCCTATT 231
DB 115 TGGGGTCTGGACATGATCTCCACACAGGACCTGTTCCGGATGAAACGTCCTATT 174
QY 232 TCAATAGGATCTTCTGCGCCACCATCTACTTTCATCATCTTCTTGACTGGCATAGTCGCA 291
DB 175 TCAATAGGATCTTCTGCGCCACCATCTACTTTCATCATCTTCTTGACTGGCATAGTCGCA 234
QY 292 ATGGATGTGATCTCTGTCATGGTTACAGAGAGTAGAGCATGACGGACAGT 351
DB 235 ATGGATGTGATCTCTGTCATGGTTACAGAGAGTAGAGCATGACGGACAGT 294
QY 352 ACCGGCTGACCTGTCACTGGGTGACCTCTCTTTGTATCATCACTCCCTCTCTGGCAG 411
DB 295 ACCGGCTGACCTGTCACTGGGTGACCTCTCTTTGTATCATCACTCCCTCTCTGGCAG 354
QY 412 TTGATGCCATGGTGTACTGGTATCTTTGGGAAATTTTGTGTAAGGTGTCATATCATCT 471
DB 355 TTGATGCCATGGTGTACTGGTATCTTTGGGAAATTTTGTGTAAGGTGTCATATCATCT 414
QY 472 ACATGTCAACCTACAGAGCTTCTCATCTGCGCTTTCATCAGCTGGACCGGTACC 531
DB 415 ACATGTCAACCTACAGAGCTTCTCATCTGCGCTTTCATCAGCTGGACCGGTACC 474
QY 532 TCGCCATTTGCCAGCCACCAACAGTCAAGGCCAAGGAAACTGTGTGTGAAAGGGCAG 591
DB 475 TCGCCATTTGCCAGCCACCAACAGTCAAGGCCAAGGAAACTGTGTGTGAAAGGGCAG 534
QY 592 TCTATGTGGCGTCTGGATCCGAGCCCTCTCTGACTATCTGACTTTCATCTTTGCGG 651
DB 535 TCTATGTGGCGTCTGGATCCGAGCCCTCTCTGACTATCTGACTTTCATCTTTGCGG 594
QY 652 ACGTGACCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACC 711
DB 595 ACGTGACCCAGGGGACATCAGTCAGGGGATGACAGGTACATCTGTGACCGCTTTACC 654
QY 712 CCGATAGCCTGTGATGGTGGTGTTCATTCAGCATATATAGT-GGGTCTCATCTG 770
DB 655 CCGATAGCCTGTGATGGTGGTGTTCATTCAGCATATATAGTGGGGGTCTCATCTG 714
QY 771 CCGGGATCGTCATCTCTCTGTTACTGTGATCATCATCTCTTAAGCTGTGACATCCAAAG 830
DB 715 CCGGGATCGTCATCTCTCTGTTACTGTGATCATCATCTCTTAAGCTGTGACATCCAAAG 774
QY 831 GG-----CCACAGAGAGCGCCCTCAAGACGACAG 864
|||||
```

```
Db 775 GGGCCACCAGAAACGCCAGGCCCTCAAGACGACAG 811

RESULT 2
BX397685
LOCUS
DEFINITION
BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI041Y013 5-PRIME, mRNA sequence.
ACCESSION
BX397685
VERSION
BX397685.1 GI:30621359
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 1091)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI041AH07QPkcluster=1272.r. Contact :
Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI041AH07QPl.

FEATURES
Location/Qualifiers
1..1091
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI041Y013"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 38.4%; Score 720.4; DB 13; Length 1091;
Best Local Similarity 82.1%; Pred. No. 2.1e-73;
Matches 881; Conservative 2; Mismatches 169; Indels 21; Gaps 4;

QY 38 GCCCGGCGAGGTGCGAGTAGCAGTACCCTCTGAGGCGTTTGGTGTCCGGTAAACCA 97
DB 38 GCTGTACCGGTCCGGAATTCGCGGATCCCGAGGCGCTGAGTGTCCAGTAGCCACCG 97
QY 98 OGGCTGTAGAGCGAGTGTTCGATGGAACCGATCAGTGTGAGTATATACACTTCTGATA 157
DB 98 CATCTGGAGAACCACGCGTTACCAATGGA-----GGGATCAGTATATACACTTCA 152
QY 158 CTACTCTGAGAAGTGGGGTCTGGAGACTATGACTCCAAAGGAACCCCTGCTTCCGGGA 217
DB 153 CTACACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCTGTTCGGTGA 212
QY 218 TGAAGACGTCCATTTCAATAGGATCTTCTGCGCCACCATCTCTCTTCTTCTTAC 277
DB 213 AGAAATGCTAATTTCAATAAATCTTCTGCGCCACCATCTCTCTTCTTCTTAC 272
QY 278 TGGCATAGTCGCAATGGATTTGGTGTGATCCCTGCTGATGGTTACCAAGAAAGCTAAGGAG 337
DB 273 TGGCATTTGGGCAATGGATTTGGTGTGATCCCTGCTGATGGTTACCAAGAAAGCTAAGGAG 332
QY 338 CATGACGACAAAGTACCGGCTGCACTCTGTAGTGGGTGACCTCTCTTGTGTATCACT 397
DB 333 CATGACGACAAAGTACAGGCTGCACTCTGTAGTGGGCGACCTCTCTTGTGTATCACT 392
```



```

Db 732 GCCTGGTATTGTCATCTCTGCTGCTATTATGCTATATCATCTCCAGCTGTACACATCCAA 791
Qy 830 GGGCCACAGAGCGCAGAGCCCTCAGACACAGTATCTCTATCTAGCTTCTTTTC 889
Db 792 GGGCCACAGAGCGCAGAGCCCTCAGACACAGTATCTCTATCTAGCTTCTTTTCGC 851
Qy 890 CTGCTGGCTGCCATATTATGTTGGGATCAGATCGATCTCTTCTATCTTTTGGAGTCAT 949
Db 852 CTGTTGGCTGCTTACTACTATGTTGGATCAGATCGATCTCTTCTATCTCTCTGGAATCAT 911
Qy 950 CAAGCAAGGATGTCATCTCGAGAGCATTTGCAAGTGTGATCTCCATCAGAGGCGCT 1009
Db 912 CAAGCAAGGATGTCATCTCGAGAGCATTTGCAAGTGTGATTTTATCACCAGGCGCT 971
Qy 1010 CGCCTTTCTCCACTGTCCTGAAACCCCATCTCTATGCTCTCTCGGGCCAAAGTTCAA 1069
Db 972 AGCTTTTTCACCTGTTGTAACCCCTCTCTATGCTCTCTGCGGCAATTTAA 1030
Qy 1070 AAGCTCTGCCAGCATCACTCACTCAGAGCAGAGCTCCAGCTCAAGATCTCTTC 1129
Db 1031 MCTCT- --GCCASACGCMCTCACCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1083
Qy 1130 CAAGCAAGGATGTCATCTCGAGAGCATTTGCAAGTGTGATCTCCATCAGAGGCTTCA 1189
Db 1084 CAAGCAAGGATGTCATCTCGAGAGCATTTGCAAGTGTGATTTTATCTCTCTCTCTCT 1143
Qy 1190 CTCAGCTAACCTTATGCAAGATCTATATATATATATATATATATATATATATAGAA 1246
Db 1144 KTAARATTTTWTTRAAAHWTTTTWTWATTTTWTWATTTTWTWATTTTWTWATTTTWTW 1200

```

RESULT 4

LOCUS BX396365 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION clone CS0DI015YL20 5-PRIME, mRNA sequence.

ACCESSION BX396365

VERSION BX396365.1 GI:30609085

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 1201)

JOURNAL Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

COMMENT Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1272.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DI015DF10QPl&cluster=1272.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0DI015DF10QPl.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DI015YL20"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 37.6%; Score 705.6; DB 13; Length 1201;
 Best Local Similarity 82.0%; Pred. No. 9.7e-72;
 Matches 884; Conservative 7; Mismatches 161; Indels 26; Gaps 6;

```

Qy 67 TCTGAGGCGTTTGGTCTCCGGTAAACACACCGCTGTAGACGAGTGTGTCATGGAAC 126
Db 65 TCGAGGGCTGAGTGTCTCCAGTAGCCACCGCATCTGGAGAACACAGCGGTACCATGA- 123
Qy 127 CGATCAGTGTGATATATACATCTTGTATATCTCTGAAGAAGTGGGTCTGGAGACT 186
Db 124 ---GGGGATCAGTATATACATCTTCCAGATACTACACGAGGAAATGGCTCAGGGACT 179
Qy 187 ATGACTCCAAACAGGAACCCCTCTCCGGATGAAAGCTCCATTTCAATAGGATCTTCC 246
Db 180 ATGACTCCAAAGGAACCCCTCTTCCGTGAAGAAATGCTAAATTTCAATAAATCTTCC 239
Qy 247 TGCCCAACCATCTACTCTATCTCTTCTGATGCGATAGTGGCAATGATGATGATCC 306
Db 240 TGCCCAACCATCTACTCTATCTTCTTAACTGGCATTTGGGCAATGATGATGATCC 299
Qy 307 TGGTCTATGGTTACGAGAGCTTAAGGAGCATACGAGCAAGTACCGGCTGCACCTGT 366
Db 300 TGGTCTATGGTTACGAGAGCTTAAGGAGCATACGAGCAAGTACGAGCTGCACCTGT 359
Qy 367 CAGTGGCTGACCTCTCTTCTTGTATCACAACCTCCCTTCTGGGCAAGTGTGATGGGTG 426
Db 360 CAGTGGCCGACCTCTCTTGTATCACAACCTCTTCTGGGCAAGTGTGATGGGTGCAA 419
Qy 427 ACTGGTACTTTGGGAAATTTTGTGTAAGCTGTGCAATCATCTACATCTGTCAACCTCT 486
Db 420 ACTGGTACTTTGGGAACTTCTCTATGCAAGGCAAGTGTGATGATGATGATGATGAT 479
Qy 487 ACAGCAGGCTTCTCATCTCTGGGCTTCTCAGCGCTGACCGGTACCTCGCCATTTGCCAG 546
Db 480 ACAGCAGTGTCTCTCTCTGGGCTTCTCAGCTGTGACCGCTACCTGGCCATCTGCCAG 539
Qy 547 CCACCAACAGTCAAAGGCCCAAGAACTCTGCTGGCTGAAAGGAGTCTATGTGGGCTCT 606
Db 540 CCACCAACAGTCAAGAGGCCCAAGAACTCTGCTGGCTGAAAGGAGTCTATGTGGGCTCT 599
Qy 607 GGATCCAGCGCTCTCTCTGACTATATCTGACTTCTTCTGCGACGTCAGCCAGGGGG 666
Db 600 GGATCCCTGCTCTCTCTGCTGACTATTTCCGACTTCTCATCTTTGCCAAGC----- 647
Qy 667 ACATCAGTCAGGGGATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCTGTGA 726
Db 648 ---TCAGTAGGGCAGATGACAGATATATCTGTACCGCTTCTACCCCAATGATCTGTGG 704
Qy 727 TGGTGTGTTTCAATTCAGCATATATAGTGGTGTCTCATCTCTGCCGCTGCTCATCTCC 786
Db 705 TGGTGTGTTTCCAGTTTCAGCATATGTTGGCTTATCTGCTGTTATTTGTTCTATCC 764
Qy 787 TCTCCTGTTACTGCATCATCTCTTAAGCTGTACATCTCAAGGGCCACCAAGAGCGCA 846
Db 765 TGTCTCTATTGCTATTATCATCTCCAAGCTGTACACTCCAAGGGCCACCAAGAGCGCA 824
Qy 847 AGGCCCTCAAGCAGCAGTCTCTCTCATCTAGCTTTCTTGGCTGCTGGCTGCCATATT 906
Db 825 AGGCCCTCAAGCAGCAGTCTCTCTCATCTGCTGCTTTCTTGGCTGCTGGCTGCCATT 884
Qy 907 ATGTGGGGATCAGCATCGACTCTCTCATCTTTGGGAGTCTCATCAAGCAAGGATGTGACT 966
Db 885 ACATTGGGATCAGCATCGACTCTCTTCTCTCTGGAATCATCAAGCAAGGGGTGTGAGT 944
Qy 967 TCGAGAGCATTTGCGACAGTGTGATCTCATCAGAGGCTCTCGCTTCTTCCACTGTT 1026
Db 945 TTGAGAACACTGTGCAACAGTGGAT-TYCATCACCAGGCGCTTAGCTTTCTTCCACTGTT 1003
Qy 1027 GCCTGAAACCCCTCTCTATGCTCTCTCGGGGCCAAGTTCAAAGCTGTGCCAGCATG 1086
Db 1004 GTCGAA-CCATCTCTCTATGCTTCTTCTTGGCAATTTAAACTCT---GCCACAG 1059

```

[illegible]

RESULT 5	1201 bp	mRNA	linear	EST 13-MAY-2003
BX421511				
LOCUS	BX421511	Homo sapiens	B CELLS (RAMOS CELL LINE)	Homo sapiens cDNA
DEFINITION	BX421511	clone CS0DG007YM12	5-PRIME, mRNA sequence.	

EXPIRATION	
VERSION	BX421511.1 GI:30651194
EST.	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
KEYWORDS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

REFERENCE
1 (Bases 1 to 1201)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r. For
more information about this cluster, see

http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DG007BG06Q1&cluster=1272.r. Contact :
 Feng liang Email : fliang@fatech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DG007BG06Q1.
 Location/Qualifiers

```

1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C90DG007YM12"
/tissue_tore="B CELLS (RAMOS CELL LINE)"
source

```

```

/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/vector: pcwvSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcwvSPORT_6 vector.

```

```

ORIGIN
Query Match      37.5%; Score 704.6; DB 13; Length 1201;
Best Local Similarity 81.4%; Pred No. 13e-71;
Tidels          7.0e-157; 20. Gaps 1.
Library was not normalized."

```

31 AGCGGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 90
37 ACCGCTCCGGATTCCTCGGGATGTAGCAAGTAGCGCGGAGGCGCTGAGTCTCCAGTA 96
38 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 105
39 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 114
40 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 123
41 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 132
42 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 141
43 ACGGCGCGCGCGCGGCGAGTGCAGGTAGCAGTAGCACTCTGAGGCGCTTGGTCTCCGGTA 150

Db 97 GCCACCGCWTCTGGAGAACACGAGCTTACCATGGA-----GGGGATCAGATATATACACTT 151

Qy 151 CTGATTAATCTACTCTGAAGAAGCTGGGGTCTGGAGACTATGACTCCACAGGAAACCTGCT 210

Qy 211 TCGGGATGAAACGTCGATTTCAATAGGATCTCTCTGCCACCATCTACTTCTCATCT 270
 ||||| : |||||
 Db 212 TCGGTGAAGAAATGCTATTTTCAATAAATCTCTGCCACCATCTACTTCTCATCT 271
 ||||| : |||||

Db 272 TCTTAAGGCATTCTGGGCAATGGATTGGTCACTCTGGTCAATGGGTACCGAAGAMAC 331

331	TAAGGAGCATCAGCGACAAGTACCGGTGCACTGTGAGTGGCTGACCTCCTCTTTGTCA	390
332	TGAGAAAGCATAGCGACAAGTACAGCTGCACTGTGAGTGGCGCACTCCTCTTTGTCA	391
339	TCACATCCCTCTCTGGCAGTTGATGCCATGGCTGACTGTGATCTTTGGGAAAATTTTGT	450
392	TCAGCTTCCCTCTCTGGCAGTTGATGCCGTGGCAAACTGGTACTTTTGGGAATTCCTAT	451
451	GTAAGGCTGCCATATCATCTACACTGTCAACCTCTACAGCAGCGTGTCTCATCTCGGCT	510
452	GOMAGGAGTCCATGTCACTACACAGTCAACCTCTACAGCMGTGTCCTCATCTCGGCT	511
511	TCATCAGCGCTGGACCGGTACTCGGCATGTGCCAGCCACCAACAGTCAAAAGGCCAAGGA	570
512	TCATCAGTCTGGACCGCTACCTGGCCATCGTCCACGCCCAACMACAGTCAGAGGCCAAGGA	571
571	AACTCTGGCTGAAAAGGAGTCTATGTGGGGTCTGGATCCAGCCCTCCTCTGACTA	630
572	AGCTGTGGCTGAAAAGTGTCTATGTGGGCTCGGATCCCTCCCTCTCGCTGACTA	631
631	TACCTGACTTCACTTTTCCGACGTCAGCAGGGGGACATCAGTCAGGGGGATGACAGT	690
632	TTTCCGACTTCACTTTTGCACAGTCAGTGA-----NGCAGATGACAGAT	677
691	ACATCTGTGACCGCTTTACCCCGATAGCTGTGATGGTGTGTTTCAATTCAGCATA	750
678	ATATCTGTGACCGCTTCAACCCAAATGACTTTGGGGTGGTTGTGTTCCAGTTTCAGCACA	737
751	TAATGGTGGTCTCATCTCTGCCCGGCATCGTCACTCTCTCTGTTACTGCATCATCATCT	810
738	TCATGGTGGCTTATCCCTGCTGGTATGTGATCATCTGTCTCTGCTATTGCATATCATCT	797
811	CTAAGCTGTCACTCCAGGGCCACCAAGAGCGCAAGGCCCTCAAGACGACAGTCATCC	870
798	CCAGCTGTCACTCCAGGGCCACCAAGAGCGCAAGGCCCTCAAGACCAACAGTCATCC	857
871	TCATCTAGCTTTCTTTGGCTGCTGCTGCCAATATTGTGGGGATCAGCATCGATCCT	930
858	TCATCTGGCTTTCTTTGGCTGCTGGCTGCTTACTACATTGGGATCAGCATCGATCCT	917
931	TCATCTTTTGGAGTCACTAAGCAAGGATGTGACTTCGAGAGCATTTGTCACAAGTGA	990
918	TATCTCTCTGGAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTGCACAAGTGA	977
991	TTCTCATCAGAGGCCCTCGCTTTCTTCCACTGTGTGCTGAACCCCACTCTCTATGCT	1050
978	TTTTCATCAACGAGGCCCTAGTTTCTTCCACTGTGTGTGAA-CCCATCTCTATGCTT	1036

QY		1051	TCTCGGGCCAAAGTTCAAAAGCT	1074
			:	
Db		1037	TCVTTGGAGCAAATTTAAAAC TCT	1060

Q7 165 GAAGAAGTGGGGTCTGGAGACTATGACTCCACATAGGAAUCCUCCUCCCGGGAAGGAGGAC 22

ORIGIN		Tracking code 012.0	
Query Match	36.5%; Score 684.8; DB 12; Length 1051;		
Best Local Similarity	81.6%; Pred. No. 2.5e-69;		
Matches 873;	Conservative 0; Mismatches 172; Indels 25; Gaps		
54	GTAGCAGTGACCCCTGAGCGGTTGGTCTCCGGTAAACCAACCAACCGGCTGTAGACCGAGT	113	
1	GTAGCAAGTGAAGCCGAGGGTCTGACTGTCTCAGTAGCCACCGCATCTGGAGAACACGC	60	
114	GTTCGCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATAACTCTCTGAAGAAGTG	173	
61	GGTTACCATTGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGGAATG	115	
174	GGCTCTGGAGACTGACTCCACAGAGAACCTGCTCCGGGTGAAAAAGTCCATTTTC	233	
116	GGCTCAGGGACTATGACTCATGAGGAACCTGTTCCGTGAAGAAAAATGCTAAATTTTC	175	
234	AATAGGATCTTCTCGCCCAACATCTACTTATCATCATCTTCTTGACTGGCATAGTCGGCAAT	293	
176	AATAAAATCTTCTCGCCCAACATCTACTCTCATCATCTTCTTAACTGGCATTTGGGCAAT	235	
294	GGATTGGTGATCCGTGTCATGGGTACCGAGAGAGACTAAGGAGCATGACGGACAAGTAC	353	
236	GGATTGGTCATCCTGGTTCATGGGTACCGAAGAAACTTGAGAAGCTGACGGACAAGTAC	295	
354	CGGCTGCACCTGTCAAGTGGCTGACCTCCTCTTTGTTCATCACACTCCCTTCTGGGCAGTT	413	
296	AGGCTGCACCTGTCAAGTGGCGGACCTCCTCTTTGTTCATCACGCTTCCCTTCTGGGCAGTT	355	
414	GATGCCATGGCTGACTGGTACTTTGGGAATTTTGTGAAGGCTGTCCATATCATCTAC	473	
356	GAATCGGTGGCAAACTGGTACTTTGGGAATCTCTATGCAAGGACATGTCATCTAC	415	
474	ACTGTCAACCTCTACAGACGGTTCTCATCTTGGCCTTTCATCAGCCTGGACCGGTACCTC	533	
416	ACAGTCAACCTCTACAGACGTGTCCTCATCTTGGCCTTTCATCAGTCTGGACCGGTACCTG	475	

296	AGGCTGCACCTGTGACGTGGCGGCACTCCCTCTTTGTGCATCAGCGTTCGCTTGCGGCAGTT	3555
414	GATGCCATGGGTGACTGGTACTTTGGGAAATTTTGTGAAGGTGTCCTATCATCTAC	4733
356	GATGCCGTGGCAAACTGGTACTTTGGGAACTCCATGCAAGGAGTCCTATGTCATCTAC	4155
474	ATGTCAACCTCTACAGCAGCGTTCTCATCTCGGCGCTTCATCAGCGTGGACCGGTACCTC	5333
416	ACAGTCAACCTCTACAGCAGTGCCTCATCTCGGCGCTTCATCAGTCTGGACCGGTACCTG	4755

```
QY 534 GCCATTGTCCAGCCACCAACAGTCAGTCAAGAGGCGCAAGAACTGCTGGCTGAAAGGCGATC 593
Db 476 GCCATCGTCCAGCCACCAACAGTCAGTCAAGAGGCGCAAGAACTGCTGGCTGAAAGGCGATC 535
QY 594 TATGTGGGCGTCTGATCCAGCCCTCCCTCTGACTATACCTGACTGATTCATCTTTGCGGAC 653
Db 536 TATGTGGGCGTCTGATCCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
QY 654 GTCAGCCAGGGGACATCAGTCAAGGGGATGACAGGTACATCTGTGACCGCTTTACCCC 713
Db 596 G-----TCAGTGAGGAGATGACAGATATATCTGTGACCGCTTTACCCC 640
QY 714 GATAGCTGTGATGGTGGTGTTCATTCAGAGATATAAGTGGGTCTCATCTCGCC 773
Db 641 AATGACTTGTGGTGGTGTTCAGTTCAGATCATGTTGGCTTATCTCGCT 700
QY 774 GGCATGCTATCTCTCTGTACTGATCATCATCATCTTAAGCTGTGACATCCCAAGGCG 833
Db 701 GGTATTGTATCTCTCTGTATTCATTCATCTCAAGCTGTGACATCCCAAGGCG 760
QY 834 CACAGAGGCGCAAGGCGCTCAAGACAGTCACTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTG 893
Db 761 CACAGAGGCGCAAGGCGCTCAAGACAGTCACTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTG 820
QY 894 TGGCTGCCATATATGTGGGATGACATCGACTCCCTTCATCTTTTGGGATCATCAAG 953
Db 821 TGGCTGCCATATATGTGGGATGACATCGACTCCCTTCATCTTTTGGGATCATCAAG 880
QY 954 GAA-GGATGTGACTTCAGAGATTTGTGACAAAGTGGATCCCATCAGAG-GGCGCTCG 1011
Db 881 CAAGGGGTGTAGTTGTGAGAACAGCGTGCACAAAGTGGATTTTCATCAGGAGGCGCTGAG 940
QY 1012 CTTCTTCACCTGTTGCTGAAACCCATCTCTATGCTT--CCTCGGGGCGCAAGTTCAA 1069
Db 941 CTTCTTCACCTGTTGCTGAAACCCATCTCTATGCTT--CCTCGGGGCGCAAGTTCAA 1000
QY 1070 AAGCTCTGCCAGATGACTCAATCCATG-AGCAGAGGCTCCAGCTC 1118
Db 1001 AACCTCTGCCAGATGACTCAATCCATG-AGCAGAGGCTCCAGCTC 1050
```

```
RESULT 9
BX462627 1201 bp mRNA linear EST 22-MAY-2003
LOCUS BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0DH005YK1.0 5-PRIME, mRNA sequence.
ACCESSION BX462627.1 GI:31029445
VERSION 1 (bases 1 to 1201)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH005BF05P1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DH005BF05QPI.
FEATURES
source Location/Qualifiers
1..1201 /organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH005YK1.0"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed,
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
```

```
ORIGIN
Query Match 35.2%; Score 660.8; DB 13; Length 1201;
Best Local Similarity 82.3%; P-red. No. 1.2e-66;
Matches 815; Conservative 0; Mismatches 152; Indels 23; Gaps 4;

QY 55 TAGCAGTACCCCTCTGAGCGCTTGGTCTCGGTAAACACACCGCTGTAGAGGAGTG 114
Db 67 TAGCAAAAGTGACGCCGAGGGCTGAGTCTCCAGTAGCCCGCATCTGAGAACCGAG 126
QY 115 TTGCCATGAAACCGCATCAGTGTGAGTATATACATCTTGTATTAACACTCTGAAAGAGTGG 174
Db 127 GTTACCATGGA-----GGGGATCAGTATATATACATCTCAGATTAACACCGAGGAATGG 181
QY 175 GGTCTGGAGACTATGATCCAAAGGAACCTGTCTCCGGGATGAAAGCGTCCATTTC 234
Db 182 GCTCAGGGGACTATGATCCATGAAGGAACCTGTCTCCGGTGAAGAAATGCTAATTTCA 241
QY 235 ATAGGATCTTCTGCCCCACCATCTACTTCTCATCTCTTGTGACTGGCATAGTCGCAATG 294
Db 242 ATAAATCTTCTGCCCCACCATCTACTTCTCATCTCTTCTTAAGTGGCATTTGGGCAATG 301
QY 295 GATGTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
Db 302 GATGTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
QY 355 GGTCTGACCTGTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 414
Db 362 GGTCTGACCTGTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
QY 415 ATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Db 422 ATCCGCTGGCAAACTGGTACTTGGGAACTTCTATGCAAGGCGAGTCCATGTCTATCTACA 481
QY 475 CTGTCAACCTGTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 534
Db 482 CAGTCAACCTGTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
QY 535 CCATTGTCCACGCCACCAACAGTCAAGGCGCAAGGAACTGCTGGCTGAAAGGCGAGTCT 594
Db 542 CCATTGTCCACGCCACCAACAGTCAAGGCGCAAGGAACTGCTGGCTGAAAGGCGAGTCT 601
QY 595 ATGTGGGCTGTGGATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 654
Db 602 ATGTGGGCTGTGGATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
QY 655 TCAGCCAGGGGACATCAGTCAAGGCGGATGACAGGTACATCTGTGACCGCTTTTACCCCG 714
Db 662 -----TCAGTGAGGCGAGATGACAGATATATCTGTGACCGCTTTTACCCCG 706
QY 715 ATAGCTCTGTGGATGGTGGTGTTCATTCAGAGATATAATGGTGGTGTCTCATCTCGCCCG 774
Db 707 ATGACTTGTGGTGGTGTTCAGTTCAGACATCATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 766
QY 775 GCATCTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
Db 767 GTATTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826
QY 835 ACCAGAGGCGCAAGGCGCTCAAGACGACAGTCACTCTCATCTCTAGTCTTCTTCTGCTGCTGCT 894
Db 827 ACCAGAGGCGCAAGGCGCTCAAGACGACAGTCACTCTCATCTCTGCTGCTGCTGCTGCTGCTGCT 886
```

QY 895 GGCTGCGATATTATGTGGGATCAGATCGACTCCTTTCATCCTTTTGGGAGTCATCAAGC 954
 DB 887 GGCTGCGATATTATGTGGGATCAGATCGACTCCTTTCATCCTTTTGGGAGTCATCAAGC 946
 QY 955 AAGATGTGACTTCGAGAGCATTTGTGCAAAAGTGGATCTCCATCAGAGGCGCTCGCT 1014
 DB 947 AAGGGTGTGAGTTGAGAACACTGTGCACAAAGTGGAT-TTCATCACCAGGCGC--TAGCT 1003
 QY 1015 TCTTCCACTGTGCTGAAACCCCATCCTCT 1044
 DB 1004 TTCTTCACTGTGTCTGAACCCATCTCTAT 1033
 RESULT 10
 BX418530 1201 bp mRNA linear EST 13-MAY-2003
 LOCUS BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 DEFINITION CS0DF006XJ15 5-PRIME, mRNA sequence.
 ACCESSION BX418530
 VERSION BX418530.1 GI:30658606
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1272.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DF006DE08QPI&cluster=1272.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DF006DE08QPI.
 Location/Qualifiers
 1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF006XJ15"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."
 FEATURES
 source
 Query Match 34.5%; Score 647.4; DB 13; Length 1201;
 Best Local Similarity 81.7%; Pred. No. 41e-65;
 Matches 823; Conservative 4; Mismatches 155; Indels 25; Gaps 6;
 QY 39 CCGGGGAGGTGAGGTAGCAGTGAACCTCTGAGGCGTTGGTGTCCGGTAACCAACCAC 98
 DB 61 CCGGGATGAGGAGGTAGCAAAAGTACGCCGAGGCGCTGAGTGTCCAGTAGCAACCCG 120
 QY 99 GGCTGTAGAGGAGTGTGCGATCGGACCGATCAGTGTGAGTATATACACTTCTGATAAC 158
 DB 121 ATCTGGAGAACCGCGGTACCTAGGAGGA-----TCAGTATATACACTTCAGATAAC 175
 QY 159 TACTCTGAAGAAGTGGGCTGGAGACTATGACTCCAAAGGAACCCCTGCTTCGGGAT 218
 DB 176 TACACCAGGAATGGGCTCAGGAGACTATGAVTCCATGAAGGAACCCCTGTTCCGTGAA 235

QY 219 GAAACGTCATTTCAATAGGATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTGA 278
 DB 236 GAAATGCTAATTTCAATAGGATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTGA 295
 QY 279 GGCATAGTCGCGCAATGATTTGGTGTATCTTGTGATCTTGTGATGGTTACCAAGAGAGCTTAAGAGC 338
 DB 296 GGCATAGTCGCGCAATGATTTGGTGTATCTTGTGATCTTGTGATGGTTACCAAGAGAGCTTAAGAGC 355
 QY 339 ATGACGGAACAAGTACCGGCTGCACCTGTCACTGTGCTGCTGCTTCTTGTGATCAGACTC 398
 DB 356 ATGACGGAACAAGTACAGGCTGCACCTGTCACTGTGCTGCTGCTTCTTGTGATCAGACTC 415
 QY 399 CCTTCTGGGAGTGTATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
 DB 416 CCTTCTGGGAGTGTATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475
 QY 459 TCCATATCATCTACACTGTCAACTCTTACAGCAGCGTTCTTCTGCTGCTGCTGCTGCTGCT 518
 DB 476 TCCATATCATCTACAGTCAACTCTTACAGCAGTGTCTTCTGCTGCTGCTGCTGCTGCT 535
 QY 519 CTGACCGGTACCTGCGCATTTGTCAAGCCACCAACAGTCAAGGCGCAAGAAATCTGTG 578
 DB 536 CTGACCGGTACCTGCGCATTTGTCAAGCCACCAACAGTCAAGGCGCAAGAAATCTGTG 595
 QY 579 GCTGAAGAGGAGTCTATGTGGGCTCTGGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
 DB 596 GCTGAAGAGGAGTCTATGTGGGCTCTGGATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
 QY 639 TTCATCTTTGCGAGCTCAGCCAGGCGGACATAGTCAAGGCGGATGACAGGTATCTGTG 698
 DB 656 TTCATCTTTGCGAGCTCAGCCAGGCGGACATAGTCAAGGCGGATGACAGGTATCTGTG 700
 QY 699 GACGCGCTTTACCCGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
 DB 701 GACGCGCTTTACCCGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
 QY 759 GGTCTCATCTGCGCGGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
 DB 761 GGTCTCATCTGCGCGGATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
 QY 819 TCACACTCCAAAGGCGCACCAAGCGGAGCCCTCAAGAGCAGTCACTCTCATCTCT- 877
 DB 821 TCACACTCCAAAGGCGCACCAAGCGGAGCCCTCAAGAGCAGTCACTCTCATCTCT- 880
 QY 878 AGCTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 937
 DB 881 GGTCTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
 QY 938 TTGCGAGTCTCAAGCAAGGATGTGACTTCGAGAGCATTTGTGCAAGAGTGGATCTCCAT 997
 DB 941 CCT-GGAAATATCAAGCAAGGAGTGTGAGTTTGAGAACACACTGTGCAAGAGTGGAT-TCCAT 998
 QY 998 CACAGGCGCTGCGCTTCTTCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
 DB 999 CACGAGGCGC--TAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043

RESULT 11
 BX174412 958 bp mRNA linear EST 06-FEB-2001
 LOCUS BX174412 NCI_CGAP_Maml Mus musculus cDNA clone IMAGE:4457694 5',
 DEFINITION mRNA sequence.
 ACCESSION BX174412
 VERSION BX174412.1 GI:12681115
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 958)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0254 row: o column: 07
High quality sequence stop: 628.
Location/Qualifiers
1. .958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457694"
/tissue_type="tumor, biopsy sample"
/dev_stages="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI CGAP Maml"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;
Site 2: NCI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN
Query Match 33.6%; Score 631.4; DB 12; Length 958;
Best Local Similarity 96.2%; Pred No. 3.2e-63;
Matches 690; Conservative 0; Mismatches 21; Indels 5; Gaps 4;

QY 54 GTAGCAGTGCACCTCTGAGCGCTTTGGTCTCCGGTAAACCAACCGGTGTAGACGAGT 113
DB 1 GTAGCAGTGCACCTCTGAGCGCTTTGGTCTCCGGTAAACCAACCGGTGTAGACGAGT 60
QY 114 GTTGCATCGAACCGATCAGTGTAGTATACATCTCTGTAATCTCTCGAAGATG 173
DB 61 GTTGCATCGAACCGATCAGTGTAGTATACATCTCTGTAATCTCTCGAAGATG 120
QY 174 GGTCTCGAGACTATGACTCCAAACAGGAACCTGCTCCGGGATGAAACGTCATTTC 233
DB 121 GGTCTCGAGACTATGACTCCAAACAGGAACCTGCTCCGGGATGAAACGTCATTTC 180
QY 234 AATAGATCTTCTGCGCCACATCTACTTTCATCTCTTGTGACGTGGATCGGCAAT 293
DB 181 AATAGATCTTCTGCGCCACATCTACTTTCATCTCTTGTGACGTGGATCGGCAAT 240
QY 294 GGATTGTCATCTGTCATGTTGTTACCAAGAAAGCTAAGGAGCATGACGACAAATAC 353
DB 241 GGTTCGTGATCTGTCATGTTGTTACCAAGAAAGCTAAGGAGCATGACGACAAATAC 300
QY 354 CGGCTGCACCTGTCAGTGGCTGACCTCTCTTGTGATCACAATCCCTTCTGGGAGTT 413
DB 301 CGGCTGCACCTGTCAGTGGCTGACCTCTCTTGTGATCACAATCCCTTCTGGGAGTT 360
QY 414 GATGTCATGGCTGACCTGTTGTTGGAATTTTGTGTAAGGCTGTCATATCATCTAC 473
DB 361 GATGTCATGGCTGACCTGTTGTTGGAATTTTGTGTAAGGCTGTCATATCATCTAC 419
QY 474 ACTGTCAACCTCTACAGCAGCTTCTCATCTGGCCTTCATCAGCTGGAACCGTACCTC 533
DB 420 ACTGTCAACCTCTACAGCAGCTTCTCATCTGGCCTTCATCAGCTGGAACCGTACCTC 479
QY 534 GCCATTGTCCACGCCCAACAGCTCAAGGCCAAGAACTGCTGGCTGAAAGCGATC 593
DB 480 GCCATTGTCCACGCCCAACAGCTCAAGGCCAAGAACTGCTGGCTGAAAGCGATC 539
QY 594 TATGTGGGCTGTGGATCCAGGCCCTCCCTGACTATACCTGACTTCACTTTTGGCCGAC 653
DB 540 TATGTGGGCTGTGGATCCAGGCCCTCCCTGACTATACCTGACTTCACTTTTGGCCGAC 598
QY 654 GTCAGCCAGGGGACATCATAGTCAGGGGGATGACAGGTACATCTGTGACCCGCTTTACCC 713

Db 599 GTCCAGCAGGGGACATCAGTCAGGGGATGACAGTACTTCTGTGACCGGCTTACCCGA 658
QY 714 GATAGCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 770
Db 659 TA---GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711

RESULT 12
BX462182 1201 bp mRNA linear EST 22-MAY-2003
LOCUS BX462182 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0DG003YO19 5-PRIME, mRNA sequence.
ACCESSION BX462182
VERSION BX462182.1 GI:31031389
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DG003AH10QPKcluster=1272.r. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DG003AH10QPK1.

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG003YO19"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned, into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 33.4%; Score 626.6; DB 13; Length 1201;
Best Local Similarity 78.0%; Pred. No. 9.7e-63;
Matches 785; Conservative 29; Mismatches 166; Indels 27; Gaps 5;

QY 45 CAGGTGACAGTGTAGCAGTGCACCTCTGAGGCGTTTGGTGTCCGGTAAACCAACCGGCTGT 104
DB 50 CGGATCAGTGTAGCAGTGCACCTCTGAGGCGTTTGGTGTCCGGTAAACCAACCGGCTGT 109
QY 105 AGACGAGTGTGTCATGGAACCGATCAGTGTGAGTATATACACTTCTGATTAATCTCT 164
DB 110 AGAACCAAGCGGTACCATGGA-----GGGATCAGTATATACACTTCTGATTAATCTAC 164
QY 165 GAAGAAGTGGGCTCGAGACTATGACTCCAAACAGGAACCTGCTTCCGGGATGAAAC 224
DB 165 GAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCTGTTTCCGTGAGAAAT 224
QY 225 GTCCATTTCATAGGATCTTCTCCGCCCAACCATCTACTTCTATCATCTCTTCTTGAATGCGGATA 284
DB 225 GCTAATTTCAATAAATCTTCTCTGCCCAACCATCTACTTCTATCATCTCTTCTTAACTGGCAAT 284
QY 285 GTCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 344

Db AG-CTCAAGATCCTTT-CAAAGGAAACGGGGTGGACACTCTTCGTCTCCAGGAGTCA 537
 QY 1173 GAATCCTCCAGTTTTTCACATCCAGCTAACCTTATGCAAGACTTATATA---ATATATA 1228
 Db 538 GAATCCTCCAGTTTTTCACATCCAGCTAACCTTATGCAAGACTTATATAATATATA 597
 QY 1229 TATATATATGATAAAGAACTTTTTTATGTACACATTTTCCAGATATAAGAGACTGACCA 1288
 Db 598 TATATATATGATAGAGAACTTTTTTATGTACACATTTTCCAGATANTAGAGACTGACCC 657
 QY 1289 GTCCTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTATATGTTCCCTTAGTTT 1348
 Db 658 AGTCTGTACAGTTNNTGTTNT-----AATGACTGGTGGAGTTTAAIG-TCCCTTAGTTT 711
 QY 1349 TTGTGAGGTTTGACTTAATTTATATAAAATATGTTTTTTTGTGTTCATGTCA 1402
 Db 712 TTGTGAGGTTTGACTTAAGTTATATAATACCTGTNNTTGTGNTGNTCANGTGA 765

AA816049 639 bp mRNA linear EST 13-FEB-1998
 vrl4h02.r1 Barstead mouse myotubes MFLRB5 Mus musculus cDNA clone
 IMAGE:1120659 5' similar to GB:106797 PROBABLE G PROTEIN-COUPLED
 RECEPTOR LCR1 HOMOLOG (HUMAN); mRNA sequence.
 AA816049
 AA816049.1 GI:2985645
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 639)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMNI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LMLL ; contact the
 IMAGE Consortium (info@image.lml.gov) for further information.
 MGI:609995
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 510.
 Location/Qualifiers
 1..639
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C3H"
 /db_xref="taxon:10090"
 /clone="IMAGE:1120659"
 /cell_line="C2C12"
 /lab_host="DH10B"
 /clone_lib="Barstead mouse myotubes MFLRB5"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA
 was primed with a Not I - oligo(GT) primer [5',
 TGTTCAGAACTCAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
 31]; double-stranded cDNA was ligated to Eco RI adaptors
 [AATTGGATCCCTTG], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified pT73 vector.
 Library constructed by Bob Barstead. The C2C12 cell line
 (available from ATCC, catalog # CRL-1772) differentiates
 rapidly, forming contractile myotubes and producing

RESULT 14
 AA816049
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

Invitrogen. This sequence belongs to sequence cluster 1272.r. For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ015CH05Q01&cluster=1272.r>. Contact : Peng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DJ015CH05Q01.

FEATURES

Location/Qualifiers
1. .952
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YP09"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 33.1%; Score 621.2; DB 13; Length 952;
Best Local Similarity 82.5%; Pred. No. 4.7e-62;
Matches 738; Conservative 4; Mismatches 133; Indels 20; Gaps 2;
Qy 50 GCAGGTAGCAGTACCCCTCTGAGCGCTTGGTCTCCGGTAACCAACCGGTGTAGAGC 109
Db 68 GCAGGTAGCAAGTAGCGCGAGCGGCTGAGTCTCCAGTAGCCACCGCATCTGGAGAAC 127
Qy 110 GAGTGTTCCTCCATGGAACCGATCAGTGTAGTATATACATCTTGATAGTACTCTCTGAAGA 169
Db 128 CAGCGGTACCATGGA-----GGGGATCAGTATATACATCTTGATAGTACTCTCTGAAGA 182
Qy 170 AGTGGGTCTGGAGATATGACTCCAAAGAGAACCTGTCTCCGGGATGAAAGCTCCA 229
Db 183 AATGGGCTCAGGGGACTATCACTCCATGAAGAACCTGTCTCCGTGAAGAAATGCTAA 242
Qy 230 TTTCAATAGGATCTTCTGCGCCACCATCTACTCTCATCATCTTCTTCACTGGCATAGTCGG 289
Db 243 TTTCAATAGGATCTTCTGCGCCACCATCTACTCTCATCATCTTCTTCACTGGCATAGTCGG 302
Qy 290 CAATGGATTGGTGTATCTCTGGTCTAGGTTACCAAGAGAACCTAAGGAGATGACGGACAA 349
Db 303 CAATGGATTGGTGTATCTCTGGTCTAGGTTACCAAGAGAACCTAAGGAGATGACGGACAA 362
Qy 350 GTACCGGCTGACCTGTCACTGGCTGACCTCTCTTGTGTCATCACACTCCCTTCTGGGC 409
Db 363 GTACCGGCTGACCTGTCACTGGCTGACCTCTCTTGTGTCATCACACTCCCTTCTGGGC 422
Qy 410 AGTTGATGCCATGGCTGACCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 469
Db 423 AGTTGATGCCATGGCTGACCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCAT 482
Qy 470 CTACATGTCAACCTCTACAGAGCGGTTCTCATCTGGGCTTCATCAGCCTGGACCGGTA 529
Db 483 CTACAGTCAACCTCTACAGAGCGGTTCTCATCTGGGCTTCATCAGTCTGGACCGGTA 542
Qy 530 CCTCGCATTTGCCAGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAAGAAGGC 589
Db 543 CCTGGCCATTTGTCACGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAAGAAGGT 602
Qy 590 AGTCTATGTGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCACTCTTTGC 649
Db 603 GGTCTATGTGGGCTGTGGATCCCTGCTCTCTGCTGACTATTTCCGACTTTCATCTTTGC 662
Qy 650 CGAGTCAGCCAGGGGAGACATCAGTACAGGGGATGACAGGTACATCTGTGACCGCTTTA 709
Db 663 CAACG-----TCAGTANGGAGATGACAGATATCTGTGACCGCTTCTA 707
Qy 710 CCCCGATAGCTGTGGATGGTGTGTTCAATTCCAGCATATAATGGTGGTCTCATCT 769

Db 708 CCCCAATGACTTGTGGGTGGTGTGTTCAGTTTCCAGTATTCAGCACATCATGGTTGGCCTTATCT 767
Qy 770 GCCCGGATGTGTCATCTCTCTCTGTTTACTGATCATCATCTCTAAGCTGTACACTCCAA 829
Db 768 GCCTGGTATTGTGTCATCTCTCTCTGTTTACTGATCATCATCTCTAAGCTGTACACTCCAA 827
Qy 830 GGGCCACACAGAGCGCAAGGCCCTCAAGACGACAGTGCATCTCTATCTCTAGTTCCTTTGC 889
Db 828 GGGCCACACAGAGCGCAAGGCCCTCAAGACGACAGTGCATCTCTATCTCTGCTTTCTTCGC 887
Qy 890 CTGCTGGCTGCATATTATGTGGGGATCAGCATCGACTCTCTTCTATCTCTTTTGGGA 944
Db 888 CTGTTGGCTGCTTACTACTACATTGGGATCAGCATCGACTCTCTTCTATCTCTTCTGSPA 942

Search completed: May 17, 2004, 16:08:53
Job time : 5439.72 secs

OM protein - protein search, using sw model

Run on: May 17, 2004, 16:09:21 ; Search time 134 Seconds
(without alignments)
756.974 Million cell updates/sec

Title: US-09-367-052-2

Perfect score: 1867

Sequence: 1 MEPIVSIVYTSIDNYSEEVGS.....KRGHGSVSTESSESSPHSS 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1867	100.0	359	2	AAW64778 A murine
2	1867	100.0	359	2	AAW64778 A murine
3	1744.5	93.4	349	7	AD663141 Rat Prote
4	1744.5	93.4	349	7	AD663145 Rat Prote
5	1678.5	89.9	352	2	AA68812 Human mon
6	1678.5	89.9	352	2	AA68812 Human mon
7	1678.5	89.9	352	2	AA68812 Human mon
8	1678.5	89.9	352	2	AA68812 Human mon
9	1678.5	89.9	352	2	AA68812 Human mon
10	1678.5	89.9	352	2	AA68812 Human mon
11	1678.5	89.9	352	2	AA68812 Human mon
12	1678.5	89.9	352	2	AA68812 Human mon
13	1678.5	89.9	352	2	AA68812 Human mon
14	1678.5	89.9	352	2	AA68812 Human mon
15	1678.5	89.9	352	2	AA68812 Human mon
16	1678.5	89.9	352	2	AA68812 Human mon
17	1678.5	89.9	352	2	AA68812 Human mon
18	1678.5	89.9	352	2	AA68812 Human mon
19	1678.5	89.9	352	2	AA68812 Human mon
20	1678.5	89.9	352	2	AA68812 Human mon
21	1678.5	89.9	352	2	AA68812 Human mon
22	1678.5	89.9	352	2	AA68812 Human mon
23	1678.5	89.9	352	2	AA68812 Human mon
24	1678.5	89.9	352	2	AA68812 Human mon
25	1678.5	89.9	352	2	AA68812 Human mon

26	1672.5	89.6	352	4	AB56346 Non-endog
27	1672.5	89.6	352	2	AAW97362 G-protein
28	1670.5	89.5	352	6	ABU79148 Chemokine
29	1670.5	89.5	352	6	ABU79080 Human che
30	1667.5	89.3	352	2	AA277792 New plate
31	1621.5	86.9	337	2	AAW48734 Human RM3
32	1621.5	86.9	337	2	AAW48734 Human RM3
33	1621.5	86.9	337	2	AAW48734 Human RM3
34	1621.5	86.9	337	2	AAW48734 Human RM3
35	1621.5	86.9	337	2	AAW48734 Human RM3
36	1621.5	86.9	337	2	AAW48734 Human RM3
37	1621.5	86.9	337	2	AAW48734 Human RM3
38	1621.5	86.9	337	2	AAW48734 Human RM3
39	1621.5	86.9	337	2	AAW48734 Human RM3
40	1621.5	86.9	337	2	AAW48734 Human RM3
41	1621.5	86.9	337	2	AAW48734 Human RM3
42	1621.5	86.9	337	2	AAW48734 Human RM3
43	1621.5	86.9	337	2	AAW48734 Human RM3
44	1621.5	86.9	337	2	AAW48734 Human RM3
45	1621.5	86.9	337	2	AAW48734 Human RM3

ALIGNMENTS

RESULT 1

AAW64778
ID AAW64778 standard; protein; 359 AA.

XX AC AAW64778;

XX AC AAW64778;

DT 20-NOV-1998 (first entry)

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

XX A murine CXK chemokine receptor.

Best Local Similarity 100.0%; Pred. No. 9.7e-215;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPIVSIGIYTDNYSSEVGGSDYDSNKEPCFRDENVHFNRIFLPTIYFIILFTIGIVGNGL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MEPIVSIGIYTDNYSSEVGGSDYDSNKEPCFRDENVHFNRIFLPTIYFIILFTIGIVGNGL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 VILVMGQCKLRMTDKYRLHLHLSVADLLFVITLPPFAVDAMADWTYFGKFLCKAVHIITYV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 VILVMGQCKLRMTDKYRLHLHLSVADLLFVITLPPFAVDAMADWTYFGKFLCKAVHIITYV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 NLYSSVILILAFISLDRLYLAIVHATNSQRPKLLAEKAVTVGVWIPALLLTIPDFIFADVS 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 NLYSSVILILAFISLDRLYLAIVHATNSQRPKLLAEKAVTVGVWIPALLLTIPDFIFADVS 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 181 QGDISQGGDRYICDRLYPDSLMMVVFQFQHMVGLILPGIIVILSCYCIIIISKLHSHSGHQ 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 QGDISQGGDRYICDRLYPDSLMMVVFQFQHMVGLILPGIIVILSCYCIIIISKLHSHSGHQ 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 241 KRKALKTTVILILAFFACMLPYTVGISIDSFILLGVIKQCDFESIVHKWISITEALAFF 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 KRKALKTTVILILAFFACMLPYTVGISIDSFILLGVIKQCDFESIVHKWISITEALAFF 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 301 HCCLNPILYAFLGAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSESSFHSS 359
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 HCCLNPILYAFLGAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSESSFHSS 359
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
AAV39994
ID AAV39994 standard; protein; 359 AA.
XX
XX AAV39994;
XX
XX
DT 16-DEC-1999 (first entry)
XX
XX Mouse CXCR4 protein sequence.
DE
DE CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW tissue repairing agent; vascularisation.
KW
XX
XX Mus sp.
OS
XX
XX WO9948528-A1.
PN
PN
PD
PD 30-SEP-1999.
XX
XX
XX 23-MAR-1999; 99WO-JP001448.
XX
XX
XX 24-MAR-1998; 98JP-00095448.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX (KISH/) KISHIMOTO T.
XX
XX
XX Kishimoto T, Nagasawa T, Tachibana K;
PI
XX
XX WPI; 1999-591042/50.
XX N-PSDB; AAZ27611.
XX
XX
XX CXCR4-potentiating agents and methods useful for inhibiting
PT neovascularization, and treating solid cancers.
XX
XX
XX Disclosure; Page 49-50; 63pp; Japanese.
XX
XX
XX This sequence is the mouse CXCR4 protein. The invention relates to
CC remedies inhibiting neovascularisation, remedies for solid cancer,
CC remedies for diseases pathologically caused by neovascularisation and
CC tissue repairing agents containing as the active ingredient a substance
CC capable of potentiating CXCR4. Based on a finding that vascularisation is
CC inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
CC remedies inhibiting vascularisation which contain as the active
CC ingredient a substance capable of potentiating CXCR4, remedies for solid
CC cancer. Remedies for diseases pathologically caused by neovascularisation

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at [ftp.wipo.int/pub/published/pct](http://wipo.int/pub/published/pct) sequences.

Sequence 349 AA;

[illegible]

RESULT 4	
ADE83145	
ID	ADE631145 standard; protein; 349 AA.
XX	
XX	AC ADE631145;
XX	
XX	29-JAN-2004 (first entry)
DT	
XX	
XX	Rat Protein O08565, SEQ ID NO 9080.
DE	
XX	
XX	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX	
XX	Rattus norvegicus.
OS	

WO2003016475-A2.
27-FEB-2003.
14-AUG-2002; 2002WO-US025765.
14-AUG-2001; 2001US-0312147P.
01-NOV-2001; 2001US-0346382P.
26-NOV-2001; 2001US-0333347P.
(GEOH) GEN HOSPITAL CORP.
(FARB) BAYER AG.
Woolf C, D'urso D, Befort K, Costigan M;
WPI; 2003-268312/36.
GENBANK; O08565.
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a rat protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 349 AA;

Query Match	93.4%;	Score 1744.5;	DB 7;	Length 349;
Best Local Similarity	95.5%;	Pred. No. 4.7e-200;		
Matches 338;	Conservative 4;	Mismatches 7;	Indels 5;	Gaps 1;
Qy	6	VSIIYTSNYSEVSGSDYDSNKEPCFDENVHFNRIEPLTIYFIPLTIGIVNGGLVILVM	65	
		.		
Db	1	MEIITSNYSEVSGSDYDSNKEPCFDENVHFNRIEPLTIYFIPLTIGIVNGGLVILVM	60	
Qy	66	GYQKLRSMTDKYRLHLSVADLLFVITLPPWAVDAMADWYFGKPLCKAVHIIYTVNLYSS	125	
Db	61	GYQKLRSMTDKYRLHLSVADLLFVITLPPWAVDAMADWYFGKPLCKAVHIIYTVNLYSS	120	
Qy	126	VLIILAFISLDRLVLAIVHATNSQRPRKLLAEKAVVGVNVIIPALLTIPDFIFADYSQGDIS	185	
Db	121	VLIILAFISLDRLVLAIVHATNSQSARKLLAEKAVVGVNVIIPALLTIPDFIFADV-----S	175	
Qy	186	QGDRIYICDRILPYPSLMMVWFQFOHIVWGLILPGVILSCYCIISKLSSHSGKHQKRAK	245	
Db	176	QGDRIYICDRILPYPSLMMVWFQFOHIVWGLILPGVILSCYCIISKLSSHSGKHQKRAK	235	

This sequence is the human CXCR4 protein. The invention relates to remedies inhibiting neovascularisation, remedies for solid cancer, and remedies for diseases pathologically caused by neovascularisation and tissue repairing agents containing as the active ingredient a substance capable of potentiating CXCR4. Based on a finding that vascularisation is inhibited in a CXCR4 knockout mouse, it becomes possible to prepare remedies inhibiting vascularisation which contain as the active ingredient a substance capable of potentiating CXCR4, remedies for solid cancer, remedies for diseases pathologically caused by neovascularisation and tissue repairing agents containing as the active ingredient a substance capable of potentiating CXCR4. It is also possible to establish methods for treatment with the use of these remedies

PT hypoproliferative disorders, physical trauma, lesions and wounds.

Dislosure; Fig 14; 138pp; English.

This sequence represents CXCR-4 chemokine receptor-4 (CXCR-4). Chemokine receptors play an important role in the chemotaxis of T-cells and phagocytic cells to areas of inflammation. CXCR-4 is a member of the G-protein-coupled receptor family, which is involved in signal transduction, and its ligand is stromal cell derived factor-1 (SDF-1, AA52508). CXCR-4 also mediates CD4-independent infection by HIV-1. CXCR-4 has now been found to have a role in the aberrant proliferative behaviour of a number of cell types, including numerous primary tumours and derived cell lines. CXCR-4 is involved in cell transformation and tumorigenesis, particularly in brain, breast and colon tumours. It was found to be overexpressed in several brain tumour derived cell lines and primary brain tumour tissues, breast tumour tissues, colorectal adenocarcinoma, lung carcinoma and melanoma cell lines. CXCR-4 expression was required for the continuous proliferation of glioblastoma cancer cells, and inhibition of its gene function resulted in growth arrest. Conversely, overexpression resulted in enhanced and rapid cellular proliferation. CXCR-4 and SDF-1 can be used as markers for the diagnosis and prognosis of cell proliferative disorders, and antisense oligonucleotides complementary to at least a portion of an RNA transcript of a CXCR-4 gene can be used to inhibit hyperproliferation of a tumour cell. Agents that inhibit CXCR-4 function can be used for treating or preventing a disease or disorder involving cell overproliferation, e.g., brain cancer, breast cancer, colon cancer, prostate cancer and B cell lymphoma, and also premalignant conditions, benign tumours, hyperproliferative disorders, and benign dysproliferative disorders. They can also be used for treating e.g., cirrhosis of the liver, keloid formation, psoriasis, benign tumors, fibrocystic conditions and tissue hypertrophy. Compounds that promote CXCR-4 function can also be used for preventing or treating a disease or disorder involving a deficiency in cell proliferation, or treating a condition where cell proliferation would be desirable. Such diseases include degenerative disorders (e.g., Parkinson's disease, Alzheimer's disease), growth deficiencies, hypoproliferative disorders, physical trauma, lesions (e.g., those caused by ischaemia), and wounds

Sequence 352 AA;

Query Match 89.9%; Score 1678.5; DB 3; Length 352;

Best Local Similarity 90.1%; Pred. No. 4e-192;

Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

6 VSIYSDNYSEVSGDYSDNKEPCFDRNVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
4 ISYISDNYTEMGSDYSDNKEPCFDRNVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
66 GYQKLRSMYDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIYTVNLYSS 125
64 GYQKLRSMYDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIYTVNLYSS 123
126 VLILAFISLDRLVLAIVHATNSORPKLAEXAVYGVWIPALLTIPTDFADVSQGDIS 185
124 VLILAFISLDRLVLAIVHATNSORPKLAEXAVYGVWIPALLTIPTDFADVSQGDIS 178
186 QGDDRYICDRYPDSLMMVWFQFQHMVGLILPGVILSCYCIISKLSSHSGHQRKAL 245
179 EADDRYICDRYPDSLMMVWFQFQHMVGLILPGVILSCYCIISKLSSHSGHQRKAL 238
246 KTTVILILAFFACWLPYVYVIGSIDFILLGVTKQGDPEFIVHKWISTEALAFHCCLN 305
239 KTTVILILAFFACWLPYVYVIGSIDFILLGVTKQGDPEFIVHKWISTEALAFHCCLN 298
306 PILYAFGLAKFKTSQAHALTSVSRGSSLKLSKGRGGHSSVSTESSSPHSS 359
299 PILYAFGLAKFKTSQAHALTSVSRGSSLKLSKGRGGHSSVSTESSSPHSS 352

RESULT 9

AAE06690

ID AAE06690 standard; protein; 352 AA.

AAE06690;

16-OCT-2001 (first entry)

Human neuropeptide Y (NPY) Y3 receptor.

Human; neuropeptide Y; NPY; bone disease; bone mass; gene therapy; cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease; fracture; extracellular signal-regulated kinase; ERK; osteoporosis; osteopenia; bone metastasis; neurotransmitter; osteogenic; NPY Y3 receptor.

Homo sapiens.

WO200153477-A1.

26-JUL-2001.

22-JAN-2001; 2001WO-US002040.

20-JAN-2000; 2000US-00489872.

(BAYU) BAYLOR COLLEGE MEDICINE.

(AMLI/) AMLING M.

Amling M, Karsenty G, Ducey P;

WPI; 2001-488709/53.

N-PSDB; AAD12801.

Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing the level of neuropeptide Y activity in blood or cerebrospinal fluid.

Example 7; Page 85; 102pp; English.

The present invention relates to a method for treatment or prevention of bone diseases characterised by loss of bone mass, comprises administering to a mammal a compound that lowers the level of neuropeptide Y (NPY) in the serum or cerebrospinal fluid (CSF) or a compound that lowers the level of inositol phosphate (IP) or extracellular signal-regulated kinase (ERK). The method is specifically used to treat (including by gene therapy) or prevent osteoporosis, osteopenia or Paget's disease, but may also be used e.g., in cases of fractures or bone metastases. These diseases may also be diagnosed by detecting elevated NPY levels, including monitoring of treatment, assessing efficacy of compounds in clinical trials and for identifying subjects at risk. The present sequence is a human NPY Y3 receptor

Sequence 352 AA;

Query Match

89.9%; Score 1678.5; DB 4; Length 352;

Best Local Similarity 90.1%; Pred. No. 4e-192;

Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

6 VSIYSDNYSEVSGDYSDNKEPCFDRNVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
4 ISYISDNYTEMGSDYSDNKEPCFDRNVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
66 GYQKLRSMYDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIYTVNLYSS 125
64 GYQKLRSMYDKYRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIYTVNLYSS 123
126 VLILAFISLDRLVLAIVHATNSORPKLAEXAVYGVWIPALLTIPTDFADVSQGDIS 185
124 VLILAFISLDRLVLAIVHATNSORPKLAEXAVYGVWIPALLTIPTDFADVSQGDIS 178
186 QGDDRYICDRYPDSLMMVWFQFQHMVGLILPGVILSCYCIISKLSSHSGHQRKAL 245
179 EADDRYICDRYPDSLMMVWFQFQHMVGLILPGVILSCYCIISKLSSHSGHQRKAL 238
246 KTTVILILAFFACWLPYVYVIGSIDFILLGVTKQGDPEFIVHKWISTEALAFHCCLN 305

Db 239 KTTVILILAFACWLPYVIGISIDSFILLEIKQCEFENTVHKWISITEALAFHCLN 298
AAG79088
QY 306 PILYAFILGAKFKSAOHALNSMRGSSLKILSKGRGSHSVSTESSFSS 359
Db 299 PILYAFILGAKFKTSAOHALTSVRGSSLKILSKGRGSHSVSTESSFSS 352
RESULT 10
ID AAG79088 standard; protein; 352 AA.
XX AAG79088;
DT 10-DEC-2001 (first entry)
XX Amino acid sequence of human CXCR4 protein.
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection;
KW CXCR4.
XX Homo sapiens.
XX WO200164752-A2.
XX 07-SEP-2001.
XX 28-FEB-2001; 2001WO-US006322.
XX 02-MAR-2000; 2000US-00517605.
XX (UYNV) UNIV NEW YORK STATE.
XX (UYNV) UNIV NIJMEGEN.
XX Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX WPI; 2001-602565/68.
XX An antibody for the treatment or prevention of HIV-infection comprises a
XX gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
XX DC-SIGN due to concomitant conformational change.
XX Disclosure; Page 117-118; 131pp; English.
XX The specification describes an antibody which is specific for an
XX antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
XX is exposed upon gp120 binding of DC-SIGN due to concomitant
XX conformational change. DC-SIGN is a receptor that is specifically
XX expressed on dendritic cells and facilitates infection of T lymphocytes
XX with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
XX -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
XX with high affinity. The antibody of the invention inhibits the trans
XX enhancement of HIV entry into a T cell or macrophage facilitated by
XX dendritic cells. The antibody is useful to treat or prevent HIV
XX infection. The present sequence represents a human CXCR4 protein, which
XX is a translocation promoting agent that interacts with CD4. This receptor
XX functions in HIV-1 entry into cells
XX Sequence 352 AA;
Query Match 89.9%; Score 1678.5; DB 4; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNYSSEVSGDYDSNKPCFRDENVFNRIPLTYFIIFLTGIVGNGLVLVM 65
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 4 ISIYTSNYSSEVSGDYDSNKPCFRDENVFNRIPLTYFIIFLTGIVGNGLVLVM 63
66 GYQKRLRWTDKYLRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCAVHIITVNIYSS 125
64 GYQKRLRWTDKYLRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCAVHIITVNIYSS 123
126 VLIILAFISLDRYLAVHATNSQRPRKLLAEKAVTVGVWIPALLTIIPDFIPADVSQDIS 185

Db 124 VLIILAFISLDRYLAVHATNSQRPRKLLAEKAVTVGVWIPALLTIIPDFIPANV-S 178
QY 186 QGDDRYTCDRLYPDSLWVVVFOFOHIMVGLILPGIVILSCVCIISKLSHSKGHQKAKAL 245
Db 179 EADRYICDRFYFNDLWVVVFOFOHIMVGLILPGIVILSCVCIISKLSHSKGHQKAKAL 238
QY 246 KTTVILILAFACWLPYVIGISIDSFILGVIKQCGDFESIVHKWISITEALAFHCLN 305
Db 239 KTTVILILAFACWLPYVIGISIDSFILLEIKQCEFENTVHKWISITEALAFHCLN 298
QY 306 PILYAFILGAKFKSAOHALNSMRGSSLKILSKGRGSHSVSTESSFSS 359
Db 299 PILYAFILGAKFKTSAOHALTSVRGSSLKILSKGRGSHSVSTESSFSS 352
RESULT 11
AAG80123
ID AAG80123 standard; protein; 352 AA.
XX AAG80123;
DT 17-JAN-2002 (first entry)
XX Human CXCR4 protein.
KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
KW chronic bowel inflammation; rheumatoid arthritis; cystostatic;
KW antiinflammatory; antiasthmatic; immunosuppressive; dermatological;
KW antirheumatic; antiarthritic.
XX Homo sapiens.
XX WO200172830-A2.
XX 04-OCT-2001.
XX 02-APR-2001; 2001WO-EP003708.
XX 31-MAR-2000; 2000DE-01016013.
XX (IPFP-) IPF PHARM GMBH.
XX (FORS-) FORSSMANN U.
XX Forssmann W, Adermann K, Heitland A, Spodsberg N;
XX WPI; 2001-626256/72.
XX Diagnostic agent containing two or more receptor-specific ligands, useful
XX for detecting tumors, inflammation etc., also therapeutic use of ligand
XX inhibitors.
XX Disclosure; Page 12; 26pp; German.
XX This invention describes a novel diagnostic agent (A) comprising at least
XX two different ligands (I) for receptors (II) that are implicated in
XX disease. (A) are used for the diagnosis of tumors (especially colorectal
XX or prostatic), organ rejection, inflammation and autoimmune diseases.
XX Also inhibitors of (I) are used therapeutically against tumors (and their
XX metastases), inflammation (particularly bronchial asthma or chronic bowel
XX inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
XX where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
XX endocrine, motor or urogenital systems or skin are affected, and bone
XX marrow diseases. The products of the invention are chemokine derivatives
XX which have cytostatic, antiinflammatory, antiasthmatic,
XX immunosuppressive, dermatological, antiarthritic, antiarthritic.
XX Chemokines act on specific tumor and inflammatory cells through a
XX constellation of chemokine receptors (CR), which control migration and
XX proliferation of these cells. AAG80045-AAG80128 represent human chemokine
XX fragments used to illustrate the method of the invention
XX Sequence 352 AA;

```
Query Match      89.9%; Score 1678.5; DB 4; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGI VGNGLVILVM 65
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 ISIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGI VGNGLVILVM 63
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 66 GYQKKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCAVHIITVNLVSS 125
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 GYQKKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCAVHIITVNLVSS 123
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 126 VLI LAFISLDRLAIVHATNSQRPRLAERKAVYGVVIMPALLTIPDFIPADVSQGDIS 185
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 VLI LAFISLDRLAIVHATNSQRPRLAERKAVYGVVIMPALLTIPDFIPANV-----S 178
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 186 QGDDRYICDRYPDSLWVWFQFQIHVGLLPGVILSCYCIILSKLSHGKHQKAKAL 245
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 EADDRYICDRFPDNLWVWFQFQIHVGLLPGVILSCYCIILSKLSHGKHQKAKAL 238
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 246 KITVILILAFFACMLPYVVGISIDSFILLGVIKQCDPESIVHKWISITEALAFHCCLN 305
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 KITVILILAFFACMLPYVVGISIDSFILLGVIKQCDPESIVHKWISITEALAFHCCLN 298
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 306 PILYAFILGAKPKSSAOHALNSRGSLLKLSKRGHSHSVSTESSESSFHSS 359
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 PILYAFILGAKPKTSQAHALTSVRSGLSKLSKRGHSHSVSTESSESSFHSS 352
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
AA014003
ID AA014003 standard; protein; 352 AA.
XX AC AA014003;
XX DT 08-MAY-2002 (first entry)
XX DE Human cysteine-X-cysteine chemokine receptor 4.
XX KW Cysteine-X-Cysteine Chemokine Receptor 4; CXCR4 binding peptide;
KW Chemottractant cytokine; human immunodeficiency virus; HIV;
KW HIV envelope glycoprotein; cellular CD4 receptor; env gene; gp160; gp120;
KW gp41; CD4+ target cell; HIV-infected cell; CXCR4-binding compound;
KW acquired immunodeficiency syndrome; AIDS; human; receptor.
XX OS Homo sapiens.
XX FH Key
XX Region 1..39 Location/Qualifiers
FT /note= "Extracellular N-terminus; this is involved in HIV
FT entry into a cell"
FT 40..64
FT /note= "Transmembrane segment 1 (TM1)"
FT 65..76
FT /note= "Intracellular loop 1"
FT 77..99
FT /note= "Transmembrane segment 2 (TM2)"
FT 100..110
FT /note= "Intracellular loop 1"
FT 111..131
FT /note= "Transmembrane segment 3 (TM3)"
FT 132..176
FT /note= "Intracellular loop 2"
FT 177..197
FT /note= "Transmembrane segment 4 (TM4)"
FT 198..203
FT /note= "Extracellular loop 2; required for HIV entry into
FT a cell"
FT 204..223
FT /note= "Transmembrane segment 5 (TM5)"
FT 224..240
FT /note= "Intracellular loop 3"
FT
```

```
Region 241..261
FT /note= "Transmembrane segment 6 (TM6)"
FT 262..282
FT /note= "Extracellular loop 3; this is involved in HIV
FT entry into a cell"
FT 275
FT /note= "Binds HIV gp120"
FT 278
FT /note= "Binds HIV gp120"
FT Misc-difference 280..281
FT /note= "Apparently encoded by nucleotides 838-903 of
FT AAK98241"
FT 283..307
FT /note= "Transmembrane segment 7 (TM7)"
FT 308..352
FT /note= "Intracellular C-terminus"
XX WO200170768-A2.
XX 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US009160.
XX 21-MAR-2000; 2000US-0190946P.
XX 21-MAR-2000; 2000US-0190996P.
XX 21-MAR-2000; 2000US-0191299P.
XX 20-MAR-2001; 2001US-00813448.
XX 20-MAR-2001; 2001US-00813651.
XX 20-MAR-2001; 2001US-00813653.
XX (CONS-) CONSENSUS PHARM INC.
XX Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA;
XX WPI; 2002-010593/01.
XX N-PSDB; AAK98241.
XX Identifying agents that bind to CXCR4 chemokine receptor 4, useful for
XX treatment and prevention of acquired immunodeficiency syndrome, comprises
XX screening compound library for members that bind to the receptor.
XX Example 3; Fig 3B; 46pp; English.
XX The invention relates to a method of identifying Cysteine-X-Cysteine
XX chemokine receptor 4 (CXCR4) binding compounds. The method comprises
XX testing and selecting compounds in a library for their ability to bind to
XX CXCR4. Chemokines (chemoattractant cytokines) regulate the movement and
XX biological activities of leukocytes in many disease situations. CXCR4
XX acts as a co-receptor for human immunodeficiency virus (HIV), as it
XX interacts with the HIV envelope glycoproteins and the cellular CD4
XX receptor, to facilitate HIV entry into cells. The HIV glycoproteins
XX (encoded by the env gene) are translated as a precursor (gp160) which is
XX subsequently cleaved into gp120 and gp41. Gp120 binds to CXCR4 and the
XX CD4 receptor present on the surface of susceptible CD4+ target cells,
XX causing the virus to fuse with the cell membranes and thereby
XX facilitating entry of the virus into the cell. The eventual expression of
XX the env product on the surface of the HIV-infected cell enables the cell
XX to fuse with uninfected CD4+ cells, thereby spreading the infection. The
XX CXCR4-binding compounds of the invention inhibit the entry of HIV into
XX the CD4+ cell and therefore are useful in the prevention of HIV infection
XX and the treatment and prevention of acquired immunodeficiency syndrome
XX (AIDS). The present sequence represents the human CXCR4 protein used in
XX the method of the invention
XX Sequence 352 AA;
SQ
Query Match      89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGI VGNGLVILVM 65
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 ISIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGI VGNGLVILVM 63
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

CC anti-HIV activities. A partially purified CXCR4 protein is useful for
CC carrying out selection, identification and improvement of tight binding
CC ligands in identifying therapeutically useful compounds. Therapeutic
CC peptides, peptidomimetics, or small molecule antagonists or agonists of
CC CXCR4 binding may be used in the prevention and treatment of AIDS and HIV
CC infection. The present sequence represents human CXCR4 which is given in
CC an example from the present invention
XX
SQ Sequence 352 AA;

Query Match 89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 66 GYOKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIYTVNLVSS 125
DB 64 GYOKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIYTVNLVSS 123

QY 126 VLILAFISLDRLYLAI VHAATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVSSQGDIS 185
DB 124 VLILAFISLDRLYLAI VHAATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVSSQGDIS 178

QY 246 KTTVILILAFACWLPYVIGISIDSFILLGVIRKQGDFFESIVHKWISITEALAFPHCCLN 305
DB 239 KTTVILILAFACWLPYVIGISIDSFILLGVIRKQGDFFESIVHKWISITEALAFPHCCLN 298

QY 306 PILYAFILGAKFKSAQAHALNSMRGSSLKILSKGRGHSSVSTESSESSFFSS 359
DB 299 PILYAFILGAKFKSAQAHALNSMRGSSLKILSKGRGHSSVSTESSESSFFSS 352

RESULT 13
ABP52651
ID ABP52651 standard; protein; 352 AA.
XX
AC ABP52651;
XX
DT 30-OCT-2002 (first entry)
XX
DE Human CXCR4 chemokine receptor 4 protein SEQ ID NO:126.
XX
KW Cysteine-X-cysteine chemokine receptor 4; CXCR4 chemokine receptor 4;
KW binding; identification; AIDS; HIV; CXCR4; virucide; anti-HIV.
XX
OS Homo sapiens.
XX
EN W0200257313-A2.
XX
PD 25-JUL-2002.
XX
PF 26-OCT-2001; 2001WO-08051165.
XX
PR 27-OCT-2000; 2000US-0243587P.
PR 20-MAR-2001; 2001US-0813448.
PR 20-MAR-2001; 2001US-0813651.
PR 20-MAR-2001; 2001US-0813653.
XX
PA (CONS-). CONSENSUS PHARM INC.
XX
PI Nestor JJ, Wilson CJ, Tan Hehir CA, Kates SA, Krstenansky J;
XX
DR WPI; 2002-606378/65.
DR N-PSDB; ABQ74938.
XX
PT Identifying a binding compound for cysteine-X-cysteine chemokine receptor
PT 4 (CXCR4) for preventing or treating AIDS and HIV infection, comprises
PT binding a molecule to a molecule with a binding property corresponding to
PT CXCR4.
XX
PS Example 3; Fig 3A-B; 92pp; English.
XX
CC The present invention describes a method (M1) for identifying a binding
CC compound for cysteine-X-cysteine chemokine receptor 4 (CXCR4) comprises
CC binding a molecule from the library of two or more molecules to the
CC molecule with a binding property corresponding to CXCR4. (M1) comprises:
CC (a) providing a library of two or more molecules; (b) providing a
CC molecule with a binding property corresponding to CXCR4; (c) binding a
CC molecule from the library of two or more molecules to the molecule with a
CC binding property corresponding to CXCR4; (d) separating the bound
CC molecule from the library of two or more molecules; and (e) identifying
CC the bound molecule as a binding compound for CXCR4. Also described is a
CC binding compound (I) for CXCR4 identified from M1. (I) has virucide and

CC anti-HIV activities. A partially purified CXCR4 protein is useful for
CC carrying out selection, identification and improvement of tight binding
CC ligands in identifying therapeutically useful compounds. Therapeutic
CC peptides, peptidomimetics, or small molecule antagonists or agonists of
CC CXCR4 binding may be used in the prevention and treatment of AIDS and HIV
CC infection. The present sequence represents human CXCR4 which is given in
CC an example from the present invention
XX
SQ Sequence 352 AA;

Query Match 89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 66 VSIYTSNDYSEVSGDYDSNKEPCFRDENVHNRIPLPTIYFIIFLTGIVGNGLVILVM 65
DB 4 ISIYTSNDYSEVSGDYDSNKEPCFRDENVHNRIPLPTIYFIIFLTGIVGNGLVILVM 63

QY 66 GYOKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIYTVNLVSS 125
DB 64 GYOKLRSMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIYTVNLVSS 123

QY 126 VLILAFISLDRLYLAI VHAATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVSSQGDIS 185
DB 124 VLILAFISLDRLYLAI VHAATNSQRPRKLLAEKAVVGVWIPALLLTIPDFIFADVSSQGDIS 178

QY 186 QGDDRYICDRLYPDSLWVWFQFHIMVGLTLPGLVILSCYCIISKLSHSKGHQKRAL 245
DB 179 EADDRIYICDRLYPDSLWVWFQFHIMVGLTLPGLVILSCYCIISKLSHSKGHQKRAL 238

QY 246 KTTVILILAFACWLPYVIGISIDSFILLGVIRKQGDFFESIVHKWISITEALAFPHCCLN 305
DB 239 KTTVILILAFACWLPYVIGISIDSFILLGVIRKQGDFFESIVHKWISITEALAFPHCCLN 298

QY 306 PILYAFILGAKFKSAQAHALNSMRGSSLKILSKGRGHSSVSTESSESSFFSS 359
DB 299 PILYAFILGAKFKSAQAHALNSMRGSSLKILSKGRGHSSVSTESSESSFFSS 352

RESULT 14
ABG32977
ID ABG32977 standard; protein; 352 AA.
XX
AC ABG32977;
XX
DT 02-DEC-2002 (first entry)
XX
DE Human CXCR4 chemokine receptor 4 (CXCR4).
XX
KW CXCR4 chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; dyskinesia; Huntington's disease; restenosis;
KW inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW congestive heart failure; cardiac remodeling; angio-genic diseases;
KW solid tumour; Kaposi Sarcoma; human.
XX
OS Homo sapiens.
XX
PN US2002107195-A1.
XX
PD 08-AUG-2002.
XX
PF 17-SEP-2001; 2001US-00953692.
XX
XX 21-JUL-1998; 98US-0093596P.
PR

PR 21-JUL-1999; 99US-00358624.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Gupta SK;
XX WPI; 2002-697879/75.
DR N-PSDB; ABS53992.
XX
PT Inducing chemotaxis of endothelial cells, useful for regulating
PT angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT cell-derived factor 1 alpha.
XX
PS Claim 3; Fig 1; 26pp; English.
XX
CC The invention describes a method of inducing chemotaxis of endothelial
CC cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
CC method is used for stimulating EC chemotaxis, and thus angiogenesis, and
CC is used for treating atherosclerosis, restenosis, ischaemic stroke and
CC spinal cord injury. Inhibition of this process is useful in treatment and
CC prevention of a very wide range of diseases, such as, infection (by
CC bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
CC (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
CC benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
CC asthma, Parkinson's disease, acute heart failure, hypertension,
CC hypertension, urinary retention, osteoporosis, anginal pectoris,
CC myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
CC migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
CC schizophrenia) and dyskinesias (e.g. Huntington's disease), inflammatory
CC diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
CC disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
CC heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
CC and Kaposi Sarcoma. This is the amino acid sequence of a CXC chemokine
CC receptor 4 (CXCR4) polynucleotide
XX
SQ Sequence 352 AA;
Query Match 89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNDYSEVSGDYSNKEPCFRDENVHNFRIPLTIYFIPLTGVNGLVILVM 65
DB 4 ISYITSDNYTEBMGSDYDSMKPCFREANFNKIFLTIYSIIFLTGVNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIYTNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIYTNLYSS 123
QY 126 VLILAFISLDRLYLAIHVAHATNSORPKLLAEKAVYGVWIPALLTIPDFIADVSGDIS 185
DB 124 VLILAFISLDRLYLAIHVAHATNSORPKLLAEKAVYGVWIPALLTIPDFIADVSGDIS 178
QY 186 QGDDRYICDRLYPDSLWVWFQFHIMVGLLPGVILSCYCIILSKLSHGKHOKKAL 245
DB 179 EADDRYICDRFPNDLWVWFQFHIMVGLLPGVILSCYCIILSKLSHGKHOKKAL 238
QY 246 KTVILILAFFACWLPYYVIGSIDSFILLGVIKQCFESIVHKWISITEALAFHCCLN 305
DB 239 KTVILILAFFACWLPYYVIGSIDSFILLGVIKQCFEPENTVHKWISITEALAFHCCLN 298
QY 306 PLYAFLGAKFKSSQAHNSMRGSSILSKRGHSHSVSTESSSSPHSS 359
DB 299 PLYAFLGAKFKTSQAHTSVSRGSSILSKRGHSHSVSTESSSSPHSS 352
RESULT 15
ABG33065
ID ABG33065 standard; protein; 352 AA.
XX
AC ABG33065;
XX
DT 28-NOV-2002 (first entry)

XX Human CXCR4 protein.
XX Chemotaxis; endothelial cell; EC; angiogenesis; atherosclerosis;
KW restenosis; ischaemic stroke; spinal cord injury; infection; bacteria;
KW fungi; protozoa; virus; pain; cancer; benign prostatic hypertrophy;
KW diabetes; obesity; anorexia; bulimia; asthma; allergy; hypertension;
KW Parkinson's disease; acute heart failure; hypotension; urinary retention;
KW osteoporosis; angina pectoris; myocardial infarction; stroke; dyskinesia;
KW migraine; vomiting; psychotic disorder; neurological disorder; ulcer;
KW inflammatory bowel disease; rheumatoid arthritis; diabetic retinopathy;
KW inflammatory disease; atherosclerosis; restenosis; Kaposi sarcoma;
KW Alzheimer's disease; congestive heart failure; cardiac remodeling;
KW angiogenic disease; solid tumour; human; stromal cell derived factor-1a;
KW SDF-1alpha; CXCR4.
XX
XX Homo sapiens.
XX US2002107196-A1.
XX 08-AUG-2002.
XX 17-SEP-2001; 2001US-00953717.
XX 21-JUL-1998; 98US-0093596P.
XX 21-JUL-1999; 99US-00358624.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Gupta SK;
XX WPI; 2002-706230/76.
XX N-PSDB; ABS53750.
XX
PT Inducing chemotaxis of endothelial cells, useful for regulating
PT angiogenesis, e.g. for treating cancer, comprises treatment with stromal
PT cell-derived factor 1 alpha.
XX
PS Claim 3; Fig 1; 26pp; English.
XX
CC The present invention relates to a new method for inducing chemotaxis of
CC endothelial cells (EC). The method of the invention involves treatment
CC with a stromal cell-derived factor 1a (SDF1alpha). The method is used for
CC stimulating EC chemotaxis, and thus angiogenesis, and is used for
CC treating atherosclerosis, restenosis, ischaemic stroke and spinal cord
CC injury, while inhibition of this process is useful in treatment and
CC prevention of a very wide range of diseases, e.g. infection (by bacteria,
CC fungi, protozoa or viruses), pain, cancer and benign prostatic
CC hypertrophy, diabetes, obesity, anorexia, bulimia, asthma, allergies,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC stroke, ulcers, migraine, vomiting, psychotic and neurological disorders
CC and dyskinesias, inflammatory disorders, rheumatoid arthritis, diabetic
CC retinopathy, inflammatory bowel disease, atherosclerosis, restenosis,
CC Alzheimer's disease, congestive heart failure, cardiac remodeling,
CC angiogenic diseases, solid tumours, and Kaposi sarcoma. The present amino
CC acid sequence represents the human CXCR4 protein of the invention
XX
SQ Sequence 352 AA;
Query Match 89.9%; Score 1678.5; DB 5; Length 352;
Best Local Similarity 90.1%; Pred. No. 4e-192;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNDYSEVSGDYSNKEPCFRDENVHNFRIPLTIYFIPLTGVNGLVILVM 65
DB 4 ISYITSDNYTEBMGSDYDSMKPCFREANFNKIFLTIYSIIFLTGVNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIYTNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIYTNLYSS 123
QY 126 VLILAFISLDRLYLAIHVAHATNSORPKLLAEKAVYGVWIPALLTIPDFIADVSGDIS 185

Db	124	VLILAFISLDRYLAI	VHATNSQRPRKLLAEKVYVGVWIPALLLTIPDFIFANV-----S	178
Qy	186	QGDRIYICDELYPDSL	WMVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL	245
Db	179	EADRIYICDRFYENDL	WMVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL	238
Qy	246	KTTVILILAFACWLPY	YVIGISIDSFILLCVIRKQGDFFESIVHKWISITEALAFPHCCIN	305
Db	239	KTTVILILAFACWLPY	YVIGISIDSFILLEIIRKQCEFFENTVHKWISITEALAFPHCCIN	298
Qy	306	PILYAFILGAKFKSSA	QHALNSMGRGSSLKILSKRGKGHSSVSTESSESSFHSS	359
Db	299	PILYAFILGAKFKTSA	QHALTSVSRGSSLKILSKRGKGHSSVSTESSESSFHSS	352

Search completed: May 17, 2004, 22:16:14
 JOB time : 141 secs

OM protein - protein search, using sw model

Run on: May 17, 2004, 22:14:28 ; Search time 31 Seconds
(without alignments)
597.862 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MPEISVITSDNYSEVGS.....KRGHSSVSTSESSSFHSS 359
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1678.5	89.9	352	1 US-08-202-056-3	Sequence 3, Appli
2	1678.5	89.9	352	1 US-08-076-093A-4	Sequence 4, Appli
3	1678.5	89.9	352	1 US-08-701-265-4	Sequence 4, Appli
4	1678.5	89.9	352	2 US-08-284-586-4	Sequence 4, Appli
5	1678.5	89.9	352	2 US-08-805-478-4	Sequence 4, Appli
6	1678.5	89.9	352	2 US-08-802-627A-4	Sequence 4, Appli
7	1678.5	89.9	352	2 US-08-801-238-4	Sequence 4, Appli
8	1678.5	89.9	352	2 US-08-801-238-4	Sequence 4, Appli
9	1678.5	89.9	352	3 US-03-104-296-4	Sequence 4, Appli
10	1678.5	89.9	352	4 US-03-527-605-4	Sequence 4, Appli
11	1641.5	87.9	352	1 US-08-450-393A-6	Sequence 6, Appli
12	1641.5	87.9	352	3 US-08-446-669-6	Sequence 6, Appli
13	1641.5	87.9	352	5 PCT-US95-00476-6	Sequence 6, Appli
14	1621.5	86.9	337	1 US-08-153-848-46	Sequence 46, Appli
15	1621.5	86.9	337	3 US-09-299-843A-46	Sequence 46, Appli
16	1621.5	86.9	337	4 US-08-088-337B-46	Sequence 46, Appli
17	1621.5	86.9	337	5 PCT-US93-11153-46	Sequence 46, Appli
18	584	31.3	368	3 US-08-703-838-2	Sequence 2, Appli
19	584	31.3	368	3 US-08-829-839-2	Sequence 2, Appli
20	584	31.3	368	4 US-08-170-496D-20	Sequence 20, Appli
21	584	31.3	368	4 US-09-624-594-2	Sequence 2, Appli
22	580	31.1	368	4 US-09-170-496D-174	Sequence 174, App
23	570.5	30.6	360	3 US-08-875-573-20	Sequence 20, Appli
24	570.5	30.6	360	3 US-08-232-878-2	Sequence 2, Appli
25	570.5	30.6	360	3 US-08-045-583-55	Sequence 55, Appli
26	570.5	30.6	360	4 US-03-534-185-55	Sequence 55, Appli
27	570.5	30.6	360	4 US-08-939-107-34	Sequence 34, Appli

28	565.5	30.3	123	3 US-08-513-974B-330	Sequence 330, Appl
29	558.5	29.9	360	4 US-08-833-752-10	Sequence 10, Appl
30	550.5	29.5	378	3 US-09-239-843A-66	Sequence 66, Appl
31	550.5	29.5	378	4 US-09-088-337B-66	Sequence 66, Appl
32	544.5	29.2	355	1 US-07-759-568-1	Sequence 1, Appli
33	544.5	29.2	355	1 US-08-450-393A-8	Sequence 8, Appli
34	544.5	29.2	355	2 US-08-390-000A-5	Sequence 5, Appli
35	544.5	29.2	355	3 US-08-446-669-8	Sequence 8, Appli
36	544.5	29.2	355	5 PCT-US95-00476-8	Sequence 8, Appli
37	544.5	29.2	360	1 US-08-232-856-7	Sequence 7, Appli
38	544.5	29.2	360	4 US-09-409-778-4	Sequence 4, Appli
39	537.5	28.8	358	1 US-08-153-848-19	Sequence 19, Appli
40	537.5	28.8	358	3 US-09-299-843A-19	Sequence 19, Appli
41	537.5	28.8	358	4 US-09-088-337B-19	Sequence 19, Appli
42	537.5	28.8	358	5 PCT-US93-11153-19	Sequence 19, Appli
43	537.5	28.8	378	1 US-08-153-848-15	Sequence 15, Appli
44	537.5	28.8	378	3 US-09-299-843A-15	Sequence 15, Appli
45	537.5	28.8	378	3 US-09-251-545-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-202-056-3
; Sequence 3, Application US/08202056
; Patent No. 5440021
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,056
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-202-056-3

Query Match 89.9%; Score 1678.5; DB 1; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Db 4 ISIYTSNDYTEEMSGDYDSNKEPCFRDENANFNKIFLPTIYSIFLTGIVGNGLVILVM 63
QY 66 GYQKLRSMTDKYLRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLVYS 125
Db 64 GYQKLRSMTDKYLRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLVYS 123
QY 126 VLIILAFISLDRLYLAIIVHATNSORPRKLLAEKAVVGVWIPALLTIPDFIFADVSQGDIS 185
Db 124 VLIILAFISLDRLYLAIIVHATNSORPRKLLAEKAVVGVWIPALLTIPDFIFANV-----S 178
QY 186 QGDDRYICDRYPDLWVVPFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 245
Db 179 EADDRYICDRYPDLWVVPFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 238
QY 246 KTTVILILAFACWLPYVVGISIDSFILGVIKOGCDFESIVHKWISITEALAFPHCCLN 305
Db 239 KTTVILILAFACWLPYVVGISIDSFILGVIKOGCDFESIVHKWISITEALAFPHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALNSMRGSSILKILSKGRGCHSSVSTESSESSPHSS 359
Db 299 PILYAFILGAKFKTSQAHALTSVSRGSSILKILSKGRGCHSSVSTESSESSPHSS 352

RESULT 2
US-08-076-093A-4
; Sequence 4, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PP4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,093A
; FILING DATE: 11-Jun-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-076-093A-4

Query Match 89.9%; Score 1678.5; DB 1; Length 352;

Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNDYSEEVSGDYDSNKEPCFRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
Db 4 ISIYTSNDYTEEMSGDYDSNKEPCFRDENANFNKIFLPTIYSIFLTGIVGNGLVILVM 63
QY 66 GYQKLRSMTDKYLRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLVYS 125
Db 64 GYQKLRSMTDKYLRLHLSVADLLFVITLPPFWAVDAMADWYFGKFLCKAVHIIYTVNLVYS 123
QY 126 VLIILAFISLDRLYLAIIVHATNSORPRKLLAEKAVVGVWIPALLTIPDFIFADVSQGDIS 185
Db 124 VLIILAFISLDRLYLAIIVHATNSORPRKLLAEKAVVGVWIPALLTIPDFIFANV-----S 178
QY 186 QGDDRYICDRYPDLWVVPFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 245
Db 179 EADDRYICDRYPDLWVVPFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 238
QY 246 KTTVILILAFACWLPYVVGISIDSFILGVIKOGCDFESIVHKWISITEALAFPHCCLN 305
Db 239 KTTVILILAFACWLPYVVGISIDSFILGVIKOGCDFESIVHKWISITEALAFPHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALNSMRGSSILKILSKGRGCHSSVSTESSESSPHSS 359
Db 299 PILYAFILGAKFKTSQAHALTSVSRGSSILKILSKGRGCHSSVSTESSESSPHSS 352

RESULT 3
US-08-701-265-4
; Sequence 4, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PP4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,265
; FILING DATE: 22-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-Jun-1993
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-284-586-4

Query Match 89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSVEGSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
DB 4 ISIVTSNYSVEGSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
QY 66 GYQKLSMTDKYRLHLSVADLLFVITLPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
DB 64 GYQKLSMTDKYRLHLSVADLLFVITLPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 123
QY 126 VLIILAFISLDRLYLAIHATNSQRPKLLAEKAVVGVWIPALLTITPDIIFADVSQGDIS 185
DB 124 VLIILAFISLDRLYLAIHATNSQRPKLLAEKAVVGVWIPALLTITPDIIFANV-----S 178
QY 186 QGDRYICDRYLPDSLMWVVFQOHIMVGLILPGIVILSCYCIISKLSHSGHQKRAL 245
DB 179 EADRYICDRYLPDSLMWVVFQOHIMVGLILPGIVILSCYCIISKLSHSGHQKRAL 238
QY 246 KTTVILLIILAFACWLPYVGVISIDSFILLGVIKOGDFESIVHKWISITEALAFHCCLN 305
DB 239 KTTVILLIILAFACWLPYVGVISIDSFILLGVIKOGDFESIVHKWISITEALAFHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALTMSVRSGLSKILSKRGHSGHSSVSTESSSPHSS 359
DB 299 PILYAFILGAKFKTSQAHALTMSVRSGLSKILSKRGHSGHSSVSTESSSPHSS 352

RESULT 5
US-08-805-478-4
Sequence 4, Application US/08805478
Patent No. 5874543
GENERAL INFORMATION:
APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/805,478
FILING DATE: 25-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993

LENGTH: 352 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-701-265-4

Query Match 89.9%; Score 1678.5; DB 1; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSVEGSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
DB 4 ISIVTSNYSVEGSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
QY 66 GYQKLSMTDKYRLHLSVADLLFVITLPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
DB 64 GYQKLSMTDKYRLHLSVADLLFVITLPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 123
QY 126 VLIILAFISLDRLYLAIHATNSQRPKLLAEKAVVGVWIPALLTITPDIIFADVSQGDIS 185
DB 124 VLIILAFISLDRLYLAIHATNSQRPKLLAEKAVVGVWIPALLTITPDIIFANV-----S 178
QY 186 QGDRYICDRYLPDSLMWVVFQOHIMVGLILPGIVILSCYCIISKLSHSGHQKRAL 245
DB 179 EADRYICDRYLPDSLMWVVFQOHIMVGLILPGIVILSCYCIISKLSHSGHQKRAL 238
QY 246 KTTVILLIILAFACWLPYVGVISIDSFILLGVIKOGDFESIVHKWISITEALAFHCCLN 305
DB 239 KTTVILLIILAFACWLPYVGVISIDSFILLGVIKOGDFESIVHKWISITEALAFHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALTMSVRSGLSKILSKRGHSGHSSVSTESSSPHSS 359
DB 299 PILYAFILGAKFKTSQAHALTMSVRSGLSKILSKRGHSGHSSVSTESSSPHSS 352

RESULT 4
US-08-284-586-4
Sequence 4, Application US/08284586
Patent No. 5840856
GENERAL INFORMATION:
APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,586
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A
FILING DATE: 11-Jun-1993
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; TELEPHONE: 415/952-9881
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2PID2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-802-627A-4
Query Match 89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
;
QY 6 VSIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
;
DB 4 ISIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
;
QY 66 GYQKLSMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
;
DB 64 GYQKLSMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 123
;
QY 126 VLILAFSLDRYLAIHATNSQRPRKLLAEKAVVGVWIPALLTTPDFADVSQGDIS 185
;
DB 124 VLILAFSLDRYLAIHATNSQRPRKLLAEKAVVGVWIPALLTTPDFADVSQGDIS 178
;
QY 186 QGDDRYICDRYPDLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 245
;
DB 179 EADDRIICDRYPDLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 238
;
QY 246 KTTVILILAFACWLPYYVVGISIDSFILLGVKQCDPESIVHKWISITEALAFPHCLN 305
;
DB 239 KTTVILILAFACWLPYYVVGISIDSFILLGVKQCDPESIVHKWISITEALAFPHCLN 298
;
QY 306 PILYAFILGAKFKSSAQAHLNSMRGSSLKILSKRGHSHSVSTESSESSSFHSS 359
;
DB 299 PILYAFILGAKFKTSQAHLNLSMRGSSLKILSKRGHSHSVSTESSESSSFHSS 352
;
;
RESULT 7
US-08-801-238-4
; Sequence 4, Application US/08801238
; Patent No. 5919896
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802.627A
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; TELEPHONE: 415/952-9881
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-805-478-4
Query Match 89.9%; Score 1678.5; DB 2; Length 352;
Best Local Similarity 90.1%; Pred. No. 6.6e-133;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
;
QY 6 VSIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
;
DB 4 ISIYTSNYSSEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
;
QY 66 GYQKLSMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
;
DB 64 GYQKLSMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 123
;
QY 126 VLILAFSLDRYLAIHATNSQRPRKLLAEKAVVGVWIPALLTTPDFADVSQGDIS 185
;
DB 124 VLILAFSLDRYLAIHATNSQRPRKLLAEKAVVGVWIPALLTTPDFADVSQGDIS 178
;
QY 186 QGDDRYICDRYPDLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 245
;
DB 179 EADDRIICDRYPDLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 238
;
QY 246 KTTVILILAFACWLPYYVVGISIDSFILLGVKQCDPESIVHKWISITEALAFPHCLN 305
;
DB 239 KTTVILILAFACWLPYYVVGISIDSFILLGVKQCDPESIVHKWISITEALAFPHCLN 298
;
QY 306 PILYAFILGAKFKSSAQAHLNSMRGSSLKILSKRGHSHSVSTESSESSSFHSS 359
;
DB 299 PILYAFILGAKFKTSQAHLNLSMRGSSLKILSKRGHSHSVSTESSESSSFHSS 352
;
;
RESULT 6
US-08-802-627A-4
; Sequence 4, Application US/08802627A
; Patent No. 5892017
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802.627A
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

[illegible][illegible]

```
; ADDRESS: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinFatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,296
; FILING DATE: 24-June-1998
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/664228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
; US-09-104-296-4
;
; Query Match 89.9%; Score 1678.5; DB 3; Length 352;
; Best Local Similarity 90.1%; Pred. No. 6.6e-133;
; Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
;
; QY 6 VSIYTSNYSSEVGSDYDNKEPCPRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
; DB 4 ISIYTSNYSSEVGSDYDNKEPCPRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
;
; QY 66 GYOKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIITVNLVSS 125
; DB 64 GYOKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIITVNLVSS 123
;
; QY 126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLLTIPDFADVSQGDIS 185
; DB 124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLLTIPDFADVSQGDIS 178
;
; QY 186 QGDDRYICDRLYPDSLMMVVFQFHIMVGLIIPGIVILSCYCIILSKLSHGKHQRKAL 245
; DB 179 EADDRYICDRFYNDLWVVFQFHIMVGLIIPGIVILSCYCIILSKLSHGKHQRKAL 238
;
; QY 246 KTVILILAFFACWLPYVIGISIDSPILLGVKQGDPEFIVHKWISITEALAFPHCCLN 305
; DB 239 KTVILILAFFACWLPYVIGISIDSPILLGVKQGDPEFIVHKWISITEALAFPHCCLN 298
;
; QY 306 PILYAFILGAKFKSSAQAHLNMSRGSLSKILSKGKGHSHSVSTESSESSPHSS 359
; DB 299 PILYAFILGAKFKTSQAHLTSSRGSLSKILSKGKGHSHSVSTESSESSPHSS 352
;
; RESULT 11
; US-08-450-393A-6
; Sequence 6, Application US/08450393A
; Patent No. 5707815
; GENERAL INFORMATION:
; APPLICANT: Charo, Israel
; APPLICANT: Coughlin, Shaun
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
; TITLE OF INVENTION: PROTEIN RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,393A
;
; Sequence 4, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Ivette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
;
; US-09-517-605-4
;
; Query Match 89.9%; Score 1678.5; DB 4; Length 352;
; Best Local Similarity 90.1%; Pred. No. 6.6e-133;
; Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
;
; QY 6 VSIYTSNYSSEVGSDYDNKEPCPRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
; DB 4 ISIYTSNYSSEVGSDYDNKEPCPRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 63
;
; QY 66 GYOKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIITVNLVSS 125
; DB 64 GYOKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIITVNLVSS 123
;
; QY 126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLLTIPDFADVSQGDIS 185
; DB 124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLLTIPDFADVSQGDIS 178
;
; QY 186 QGDDRYICDRLYPDSLMMVVFQFHIMVGLIIPGIVILSCYCIILSKLSHGKHQRKAL 245
; DB 179 EADDRYICDRFYNDLWVVFQFHIMVGLIIPGIVILSCYCIILSKLSHGKHQRKAL 238
;
; QY 246 KTVILILAFFACWLPYVIGISIDSPILLGVKQGDPEFIVHKWISITEALAFPHCCLN 305
; DB 239 KTVILILAFFACWLPYVIGISIDSPILLGVKQGDPEFIVHKWISITEALAFPHCCLN 298
;
; QY 306 PILYAFILGAKFKSSAQAHLNMSRGSLSKILSKGKGHSHSVSTESSESSPHSS 359
; DB 299 PILYAFILGAKFKTSQAHLTSSRGSLSKILSKGKGHSHSVSTESSESSPHSS 352
;
; RESULT 10
; US-09-517-605-4
```


REGISTRATION NUMBER: 20.121
REFERENCE/DOCKET NUMBER: 5555-291
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-977-1001
TELEFAX: 310-977-1003
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US95-00476-6

Query Match 87.9%; Score 1641.5; DB 5; Length 352;
Best Local Similarity 88.7%; Pred. No. 8.2e-130;
Matches 314; Conservative 18; Mismatches 17; Indels 5; Gaps 1;

QY 6 VSIYSDNYSVEGSDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVM 65
DB 4 LSIYSDNYTEMGSDYDSNKEPCFREANFNKIFLPYIYSIIFLYGIVGNGLVILVM 63
QY 66 GYQKLRMTDKYRHLHLSVADLLFVITLPPWAVDAMADWYFGKFLCKAVHIIYTNVLYSS 125
DB 64 GYQKLRMTDKYRHLHLSVADLLFVITLPPWAVDAMADWYFGKFLCKAVHIIYTNVLYSS 123
QY 126 VLLAFISLDRLYLAIVHATNSORPRKLAEKAVVGVWIPALLTIPDFADYSQGDIS 185
DB 124 VLLAFISLDRLYLAIVHATNSORPRKLAEKAVVGVWIPALLTIPDFADYSQGDIS 178
QY 186 QGDDRYICDRYPDSLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKERKAL 245
DB 179 EADDRYICDRYPDNLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKERKAL 238
QY 246 KTVLLIILAFACWLPYVIGSIDSPILLGVIKGCDPESIVHKWISITALEAFHCHCLN 305
DB 239 KYVILLIILAFACWLPYVIGSIDSPILLGVIKGCDPESIVHKWISITALEAFHCHCLN 298
QY 306 PILYAFGLAKFKYSAQAHALNMSRGSLSKLSKGRGHSVSTESSESSSFHSS 359
DB 299 PILYAFGLAKFKYSAQAHALNMSRGSLSKLSKGRGHSVSTESSESSSFHSS 352

RESULT 14
US-08-153-848-46
Sequence 46, Application US/08153848
Patent No. 5759804

GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marsh, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,848
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452

FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-153-848-46

Query Match 86.9%; Score 1621.5; DB 1; Length 337;
Best Local Similarity 90.4%; Pred. No. 3.7e-128;
Matches 309; Conservative 17; Mismatches 11; Indels 5; Gaps 1;

QY 18 VSGGDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGLVILVMGYQKLSMTDK 77
DB 1 MSGGDYDSNKEPCFREANFNKIFLPYIYSIIFLTGIVGNGLVILVMGYQKLSMTDK 60
QY 78 YRHLHLSVADLLFVITLPPWAVDAMADWYFGKFLCKAVHIIYTNVLYSSVLLILAFISLDY 137
DB 61 YRHLHLSVADLLFVITLPPWAVDAMADWYFGKFLCKAVHIIYTNVLYSSVLLILAFISLDY 120
QY 138 LAIVHATNSORPRKLAEKAVVGVWIPALLTIPDFADYSQGDISQDDRYICDRLY 197
DB 121 LAIVHATNSORPRKLAEKAVVGVWIPALLTIPDFADYSQGDISQDDRYICDRLY 175
QY 198 POSLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKERKALKTIVILLAFPA 257
DB 176 PNDLWVVFQFHIMVGLILPGIVILSCYCIISKLSHSGHOKERKALKTIVILLAFPA 235
QY 258 CMLPYVIGSIDSPILLGVIKGCDPESIVHKWISITALEAFHCHCLNPIYAFGLAKFK 317
DB 236 CMLPYVIGSIDSPILLGVIKGCDPESIVHKWISITALEAFHCHCLNPIYAFGLAKFK 295
QY 318 SQAQAHALNMSRGSLSKLSKGRGHSVSTESSESSSFHSS 359
DB 296 TSAQAHALNMSRGSLSKLSKGRGHSVSTESSESSSFHSS 337

RESULT 15
US-09-299-843A-46
Sequence 46, Application US/09299843A
Patent No. 6107475

GENERAL INFORMATION:
APPLICANT: Godiska, Ronald
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,843A
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA: US 09/088,337
APPLICATION NUMBER: 01-JUN-1998
FILING DATE: 01-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,848
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,452
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Jill E. Uhl
REGISTRATION NUMBER: 43,213
REFERENCE/DOCKET NUMBER: 27866/32059B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX:
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-843A-46

OM protein - protein search, using sw model

Run on: May 17, 2004, 22:19:12 ; Search time 547 Seconds
(without alignments)
182.625 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MPEISVSYTSDNYSEVGS.....KRGHSSVSTSESSSFHSS 359

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUBCOMB.pep.*
- 7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678.5	89.9	352	9	US-09-953-692-2
2	1678.5	89.9	352	9	US-09-953-717-2
3	1678.5	89.9	352	9	US-09-104-063-4
4	1678.5	89.9	352	12	US-10-211-462-81
5	1678.5	89.9	352	12	US-10-666-689-4
6	1678.5	89.9	352	12	US-10-181-906-10
7	1678.5	89.9	352	12	US-09-813-651B-85
8	1678.5	89.9	352	12	US-10-151-274-4
9	1678.5	89.9	352	12	US-10-170-385-331
10	1678.5	89.9	352	14	US-10-225-567A-76
11	1678.5	89.9	352	14	US-10-245-850-1
12	1678.5	89.9	352	14	US-10-251-703-38
13	1678.5	89.9	352	14	US-10-021-660-123
14	1678.5	89.9	352	14	US-10-014-322A-126
15	1678.5	89.9	352	14	US-10-239-423-79

16	1678.5	89.9	352	15	US-10-160-401-3	Sequence 3, Appl1
17	1678.5	89.9	352	15	US-10-241-434-178	Sequence 178, Appl1
18	1678.5	89.9	352	15	US-10-372-683-2	Sequence 2, Appl1
19	1678.5	89.9	352	15	US-10-440-464-78	Sequence 78, Appl1
20	1678.5	89.9	352	15	US-10-452-015-1	Sequence 1, Appl1
21	1672.5	89.5	352	11	US-09-826-509-485	Sequence 485, Appl1
22	1670.5	89.5	352	9	US-09-870-759-37	Sequence 37, Appl1
23	1670.5	89.5	352	9	US-09-870-759-144	Sequence 144, Appl1
24	1670.5	89.5	352	10	US-09-751-708A-37	Sequence 37, Appl1
25	1670.5	89.5	352	10	US-09-751-708A-144	Sequence 144, Appl1
26	969.5	51.9	209	15	US-10-292-798-688	Sequence 688, Appl1
27	964.5	51.7	209	14	US-10-017-161-790	Sequence 790, Appl1
28	584	31.3	368	12	US-10-411-284-4	Sequence 4, Appl1
29	584	31.3	368	14	US-10-251-385-20	Sequence 20, Appl1
30	584	31.3	368	14	US-10-225-567A-74	Sequence 74, Appl1
31	584	31.3	368	14	US-10-345-680-59	Sequence 59, Appl1
32	584	31.3	368	14	US-10-251-686-2	Sequence 2, Appl1
33	584	31.3	368	14	US-10-239-423-78	Sequence 78, Appl1
34	584	31.3	368	15	US-10-295-027-752	Sequence 752, Appl1
35	584	31.3	415	12	US-10-411-284-2	Sequence 2, Appl1
36	580	31.1	368	14	US-10-251-385-174	Sequence 174, Appl1
37	575	30.8	472	14	US-10-106-698-6402	Sequence 6402, Appl1
38	570.5	30.6	360	9	US-09-837-446-2	Sequence 2, Appl1
39	570.5	30.6	360	9	US-09-796-744-17	Sequence 17, Appl1
40	570.5	30.6	360	9	US-09-764-413-20	Sequence 20, Appl1
41	570.5	30.6	360	12	US-10-039-659-16	Sequence 16, Appl1
42	570.5	30.6	360	13	US-10-120-394-20	Sequence 20, Appl1
43	570.5	30.6	360	14	US-10-225-567A-66	Sequence 66, Appl1
44	570.5	30.6	360	14	US-10-164-649-55	Sequence 55, Appl1
45	570.5	30.6	360	14	US-10-231-452-48	Sequence 48, Appl1

ALIGNMENTS

RESULT 1
US-09-953-692-2
; Sequence 2, Application US/09953692
; Patent No. US20020107195A1
; GENERAL INFORMATION:
; APPLICANT: Shalley, Gupta K.
; TITLE OF INVENTION: Method for Inducing Chemotaxis in Endothelial Cells by Administering Stromal Cell Derived Factor-1
; FILE REFERENCE: P50678C1
; CURRENT APPLICATION NUMBER: US/09/953,692
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 09/358,624
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 60/093,596
; PRIOR FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Human
US-09-953-692-2

Query Match	89.9%	Score	1678.5	DB	9	Length	352
Best Local Similarity	90.1%	Pred. No.	1.1e-145				
Matches	319	Conservative	19	Mismatches	11	Indels	5
Gaps							1
QY	6	VSIVTSDNYSEVGS	GDYDSNKEPCFRDENVHFNRIPLTIYFIIFLTGIVGNGVILVM	65			
Db	4	ISIVTSDNYTEMGSDYDSNKEPCFREANFNKIFLTIIYSIFLTGIVGNGVILVM	63				
QY	66	GYQKKLRSMTDKYLRLHLSVADLLFVITLPFWADVADADWYFGKFLCKAVHIIYTNLYSS	125				
Db	64	GYQKKLRSMTDKYLRLHLSVADLLFVITLPFWADVADAVANNYFGNFKLCAVHIIYTNLYSS	123				
QY	126	VLLILAFISLDRLVLAIVHATNSORPKLAEKAVYGVWIPALLTIPDFIADVSQGBIS	185				
Db	124	VLLILAFISLDRLVLAIVHATNSORPKLAEKAVYGVWIPALLTIPDFIADVSQGBIS	178				


```

RESULT 4
US-10-211-462-81
; Sequence 81, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 81
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-81

Query Match      89.9%; Score 1678.5; DB 12; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY      6 VSIYSDNYSEEVGSDYDSNKEPCPRDENVHFNRIFLTIIYFIPLTGIVGNGLVILYM 65
DB      4 ISIYSDNTTEMSGDYDSMEKPCPFREANFNKIFLEIIYSIIIFLTGIVGNGLVILYM 63
QY      66 GYOKLRSMTDKYLHLVSADLLFVITLFPWADAMADWVFGKFLCKAVHIIVTNLYSS 125
DB      64 GYOKLRSMTDKYLHLVSADLLFVITLFPWADAVANWYFGNFKCAVHIIVTNLYSS 123
QY      126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLITIPDFADVSGDIS 185
DB      124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLITIPDFIFANV-----S 178
QY      186 QGDYVICDRLYPDSILWVVFQFQHLWVGLILPGIVILSCYCIITISKLSHSGHOKRKAL 245
DB      179 EADDYVICDRFPNDLWVVFQFQHLWVGLILPGIVILSCYCIITISKLSHSGHOKRKAL 238
QY      246 KTTVILILAFFACWLPYYVIGISIDSPILLGLVILKQGDFFSIVHKMTISITTEALAFPHCCLN 305
DB      239 KTTVILILAFFACWLPYYVIGISIDSPILLEIHKQCEFFENTVHKMTISITTEALAFPHCCLN 298
QY      306 PIIYAFPLGAKFKXSAQHAINSVRGSSLKILSKRGHGHSSVSTSESSSPHSS 359
DB      299 PIIYAFPLGAKFKTSAQHALTSVRGSSLKILSKRGHGHSSVSTSESSSPHSS 352

RESULT 5
US-10-666-689-4
; Sequence 4, Application US/10666689
; Publication No. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; FILE REFERENCE: P0706PAC2DZC1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19

```

Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNYSSEVGSDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGIYVGNGLVILVM 65
DB 4 ISIYTSNYSSEVGSDYDSNKEPCFRDENANFNKIFLTIYFIPLTGIYVGNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGNFLCKAVHIIYTVNLYSS 123
QY 126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIADVSQGDIS 185
DB 124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIFANV-----S 178
QY 186 QGDDRYICDRYLPDLSLMMVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 245
DB 179 EADDRVICDRFYNDLWVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 238
QY 246 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCDPESIVHKWISITEALAFPHCCLN 305
DB 239 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCEPENTVHKWISITEALAFPHCCLN 298
QY 306 PILYAFPLGAKFKSSAQHALNSMRGSSSLKILSKGRGGHSSVSTESSESSPHSS 359
DB 299 PILYAFPLGAKFKTSAQHALTSVRGSSSLKILSKGRGGHSSVSTESSESSPHSS 352
RESULT 7
US-09-813-651B-85
; Sequence 85, Application US/09813651B
; Publication No. US20030018438A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Hehir, Christina
; APPLICANT: Kates, Steven
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds
; FILE REFERENCE: CNS-007
; CURRENT APPLICATION NUMBER: US/09/813,651B
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 85
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-651B-85

Query Match 89.9%; Score 1678.5; DB 12; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSSEVGSDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGIYVGNGLVILVM 65
DB 4 ISIYTSNYSSEVGSDYDSNKEPCFRDENANFNKIFLTIYFIPLTGIYVGNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGNFLCKAVHIIYTVNLYSS 123
QY 126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIADVSQGDIS 185
DB 124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIFANV-----S 178
QY 186 QGDDRYICDRYLPDLSLMMVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 245
DB 179 EADDRVICDRFYNDLWVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 238

QY 246 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCDPESIVHKWISITEALAFPHCCLN 305
DB 239 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCEPENTVHKWISITEALAFPHCCLN 298
QY 306 PILYAFPLGAKFKSSAQHALNSMRGSSSLKILSKGRGGHSSVSTESSESSPHSS 359
DB 299 PILYAFPLGAKFKTSAQHALTSVRGSSSLKILSKGRGGHSSVSTESSESSPHSS 352
RESULT 8
US-10-151-274-4
; Sequence 4, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Inno
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-274-4

Query Match 89.9%; Score 1678.5; DB 12; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY 6 VSIYTSNYSSEVGSDYDSNKEPCFRDENVHFNRIPLTIYFIPLTGIYVGNGLVILVM 65
DB 4 ISIYTSNYSSEVGSDYDSNKEPCFRDENANFNKIFLTIYFIPLTGIYVGNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGKFLCKAVHIIYTVNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLFFWAVDAMADWYFGNFLCKAVHIIYTVNLYSS 123
QY 126 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIADVSQGDIS 185
DB 124 VLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPDFIFANV-----S 178
QY 186 QGDDRYICDRYLPDLSLMMVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 245
DB 179 EADDRVICDRFYNDLWVVFQHIWGLILPGIVILSCYCIISKLSHSKGHQKRAL 238
QY 246 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCDPESIVHKWISITEALAFPHCCLN 305
DB 239 KTTVILILAFACWLPYYVIGISIDSFILLGVTKQCEPENTVHKWISITEALAFPHCCLN 298
QY 306 PILYAFPLGAKFKSSAQHALNSMRGSSSLKILSKGRGGHSSVSTESSESSPHSS 359
DB 299 PILYAFPLGAKFKTSAQHALTSVRGSSSLKILSKGRGGHSSVSTESSESSPHSS 352

RESULT 9
US-10-170-385-331
; Sequence 331, Application US/10170385
; Publication No. US2003020372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan

```

; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingeman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 331
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-170-385-331

Query Match      89.9%; Score 1678.5; DB 12; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy      6 VSIYTSNDYSEVSGDYDSNKPCPRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
Db      4 ISIYTSNDYTEEMSGDYDSMKPCPREENANFNKIFLPTIYSIIIFLTGIVGNGLVILVM 63

Qy      66 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGKFLCKAVHIYTVNLYSS 125
Db      64 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGNFLCKAVHIYTVNLYSS 123

Qy      126 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFADVSGQDIS 185
Db      124 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFANV-----S 178

Qy      186 QGDDRYICDRLYPDSLMMVVFQFHIMVGLIIPGIVILSCYCIISKLSHSGHQKRAL 245
Db      179 EADDRYICDRFPNDLWVVFQFHIMVGLIIPGIVILSCYCIISKLSHSGHQKRAL 238

Qy      246 KTTVILILAFACWLPYYIGISIDSFILLGVIKQGDFFSIVHKWISITEALAFFHCCLN 305
Db      239 KTTVILILAFACWLPYYIGISIDSFILLGVIKQGDFFSIVHKWISITEALAFFHCCLN 298

Qy      306 PILYAFILGAKFKSAQHALNSMRGSSLKILSKGRGSHSVSTESSESSPHSS 359
Db      299 PILYAFILGAKFKTSAQHALTVSRGSSLKILSKGRGSHSVSTESSESSPHSS 352

RESULT 10
US-10-225-567A-76
; Sequence 76, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn C.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-225-567A-76
```

```

Query Match      89.9%; Score 1678.5; DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy      6 VSIYTSNDYSEVSGDYDSNKPCPRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
Db      4 ISIYTSNDYTEEMSGDYDSMKPCPREENANFNKIFLPTIYSIIIFLTGIVGNGLVILVM 63

Qy      66 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGKFLCKAVHIYTVNLYSS 125
Db      64 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGNFLCKAVHIYTVNLYSS 123

Qy      126 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFADVSGQDIS 185
Db      124 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFANV-----S 178

Qy      186 QGDDRYICDRLYPDSLMMVVFQFHIMVGLIIPGIVILSCYCIISKLSHSGHQKRAL 245
Db      179 EADDRYICDRFPNDLWVVFQFHIMVGLIIPGIVILSCYCIISKLSHSGHQKRAL 238

Qy      246 KTTVILILAFACWLPYYIGISIDSFILLGVIKQGDFFSIVHKWISITEALAFFHCCLN 305
Db      239 KTTVILILAFACWLPYYIGISIDSFILLGVIKQGDFFSIVHKWISITEALAFFHCCLN 298

Qy      306 PILYAFILGAKFKSAQHALNSMRGSSLKILSKGRGSHSVSTESSESSPHSS 359
Db      299 PILYAFILGAKFKTSAQHALTVSRGSSLKILSKGRGSHSVSTESSESSPHSS 352

RESULT 11
US-10-245-850-1
; Sequence 1, Application US/10245850
; Publication No. US20030124628A1
; GENERAL INFORMATION:
; APPLICANT: Burns, Jennifer M.
; APPLICANT: Miao, Zhenhua
; APPLICANT: Wei, Zheng
; APPLICANT: Howard, Maureen C.
; APPLICANT: Premack, Brett A.
; APPLICANT: Schall, Thomas J.
; APPLICANT: Chemocentryx, Inc.
; TITLE OF INVENTION: Compositions and Methods for Detecting and Treating
; TITLE OF INVENTION: Diseases and Conditions Related to Chemokine Receptors
; FILE REFERENCE: 019934-00310US
; CURRENT APPLICATION NUMBER: US/10/245,850
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/338,100
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: CXCR4 chemokine receptor
; US-10-245-850-1

Query Match      89.9%; Score 1678.5; DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

Qy      6 VSIYTSNDYSEVSGDYDSNKPCPRDENVHFNRIFLPTIYFIIFLTGIVGNGLVILVM 65
Db      4 ISIYTSNDYTEEMSGDYDSMKPCPREENANFNKIFLPTIYSIIIFLTGIVGNGLVILVM 63

Qy      66 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGKFLCKAVHIYTVNLYSS 125
Db      64 GYQKLRSMTDKYRLHLSVADLLFVITLFFWADAMADWYFGNFLCKAVHIYTVNLYSS 123

Qy      126 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFADVSGQDIS 185
Db      124 VLILAFISLDRLYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIFANV-----S 178
```


184	VLIILAFISLDXYLAIVHATNSQRPRKLABEKVVYGVWIPDALLLTPTDPIFANV-----S	178
Dd		
186	OCDRYICDRLYPSLWMVWFQFOHIMVGLILPGIIVILSCYCIITIKLSHSKGHKRKAL	245
Qy	: : : : : :	
Dd	179 EADREYICDRYPNDLWVVFQFOHIMVGLILPGIIVILSCYCIITIKLSHSKGHKRKAL	238
246	KTTVILLIAFPACWLPPYVGISIDSFILLGVIOGCDPESIVHKMISTEALAPFHCCLN	305
Qy	: : : : : :	
Dd	239 KTTVILLIAFPACWLPPYVGISIDSFILLEIIKQGCEFTVHKMISTEALAPFHCCLN	298
306	PILYAFIGAKPKSAQAHLNMSGSSLIKSLGKGGGHSSYSTSESSPFSS	359
Qy	: : : : : :	
Dd	299 PILYAFIGAKEFTSAQRLTSVRGSSLIKLSLKGGGGHSSYSTSESSPFSS	352

RESULT 12

RESULT 12

US-10-251-703-38

US-10-251-703-38

Sequence 38, Application US/10251703

Publication No. US2003014849A1

GENERAL INFORMATION:

APPLICANT: Kullopoulos, Athan

APPLICANT: Covic, Lidija

TITLE OF INVENTION: G Protein Coupled Receptor Agonists and Antagonists and

TITLE OF INVENTION: Methods of Activating and Inhibiting G Protein Coupled

TITLE OF INVENTION: Receptors Using the Same

FILE REFERENCE: NEMC-215 CIP

CURRENT APPLICATION NUMBER: US/10/251,703

CURRENT FILING DATE: 2002-09-20

PRIOR APPLICATION NUMBER: 09/841,091

PRIOR FILING DATE: 2001-04-23

PRIOR APPLICATION NUMBER: 60/198,993

PRIOR FILING DATE: 2000-04-21

NUMBER OF SEQ ID NOS: 41

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 38

LENGTH: 352

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Pepducin

OTHER INFORMATION: Peptide Sequence

US-10-251-703-38

	Query Match	89.9%	Score 1678.5	DB 14	Length 352
	Best Local Similarity	90.1%	Pred. No. 1.1e-445		
	Matches 319	Conservative	19	Mismatches 11	Indels 5
					Gaps 1
QY	6	VSITYSDNYSEEVGSDYDSNKEPCFADENVHNFRIELPTIYIIFLTGIIVGNGLVILVM	65		
				
				
				
				
Db	4	ISITYSDNYTEEMGSDYDSNKEPCFEEENANFNKIFLPTIYIIFLTGIIVGNGLVILVM	63		
				
				
				
				
QY	66	GYOKLRSMTDKYRHLHSVADLLFVITLPFWADVADADWYFGKFLCKAVHIIYITVNLVSS	125		
				
				
				
				
Db	64	GYOKLRSMTDKYRHLHSVADLLFVITLPFWADVADANVYFGNFKLCKAVHIIYITVNLVSS	123		
				
				
				
				

Qy	126	VJLAFIPISLDRYALVHATNSQRPRKLAELAEKAVVGVWIPALLTLTTPDIFADVSGDIS	185
		:	
Db	124	VJLAFIPISLDRYALVHATNSQRPRKLAELAEKAVVGVWIPALLTLTTPDIFANV----	178
		:	
Qy	186	QGDRIYCDRLPYDSLMMVWVQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAL	245
		:	
Db	179	EADRYICDRFPNDLWVWVQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAL	238
		:	
Qy	246	KTTVILILAFPCWLPYVGVGISDISPILLGVIKGCDPESIVHKWISITAEALAFPHCCIN	305
		:	
Db	239	KTTVILILAFPCWLPYVGVGISDISPILLGVIKGCEFTVHKWISITAEALAFPHCCIN	298

Qy	306	PILYAFCAKFKSSIQHALNMSRGSSIKILSKGKRGGHSSVSTESSSSFHS	359
		: :	
D6	299	PILYAFCAKFKTSACHALTSVSRGSSIKILSGKRGGHSSVSTESSSSFHS	352

RESULT 13

```

US-10-021-660-123
; Sequence 123, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: US 20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021.660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 123
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-123

Query Match      89.9%; Score 1678.5; DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.1e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

QY      6 VSIYTSDNYSVEVSGDYDSNKECFRDENVHNFPLPTIYFIIFLTGIVGNGLVILVM 65
DB      4 ISYTSDNYTEMGSGDYDSNKECFREANFNKFIPLTIYSIIFLTGIVGNGLVILVM 63

QY      66 GYQKRLRMTDKYRLHLGSVADLLFVITLPFWADAMADYFGFKLCKAVHIITVNLVSS 125
DB      64 GYQKRLRMTDKYRLHLGSVADLLFVITLPFWADAVANWYFGNFKLCKAVHIITVNLVSS 123

QY      126 VLIILAFISLDYLAIVHATNSORPKLLAEKAVVGVWIPALLLTTPDFIFADVSGDIS 185
DB      124 VLIILAFISLDYLAIVHATNSORPKLLAEKVTVGVWIPALLLTTPDFIFANV-----S 178

QY      186 QGDDRYICDRIYPOSLLMWVPQFQIHNVGLITLPGVITLSCYCIISKLSHSGHQKRAL 245
DB      179 EADDRYICDREVPNDLWVVPQFQIHNVGLITLPGVITLSCYCIISKLSHSGHQKRAL 238

QY      246 KTTVTILIAFACWLPYVVGISIDSFILGVKIGKGFESIVHKWISITELAFPHCCLN 305
DB      239 KTTVTILIAFACWLPYVVGISIDSFILLEIKGCEFEVTVHKWISITELAFPHCCLN 298

QY      306 PILYAFLGAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTSESSSFHSS 359
DB      299 PILYAFLGAKFKTSAOHALTSVRGSSLKILSKGRGHSVSTSESSSFHSS 352

```

RESULT 14
US-10-014-322A-126
; Sequence 126, Application US/10014322A
; Publication NO. US20030167129A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, Jr., John
; APPLICANT: Wilson, Carol
; APPLICANT: Tan Behir, Christina
; APPLICANT: Kates, Steven
; APPLICANT: Krstenansky, John
; TITLE OF INVENTION: Binding Compounds and Methods for Identifying Binding Compounds
; FILE REFERENCE: CNS-008
; CURRENT APPLICATION NUMBER: US/10/014,322A
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/243,587
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 09/813,651
; PRIOR FILING DATE: 2001-03-20

QY 6 VSIYTSNDYSEVSGSDYDSNKEPCFRDENVHFNRIFLTPTIYFIIFLTGIVGNGLVILVM 65
Db 4 ISIYTSNDYSEVSGSDYDSNKEPCFRDENVHFNRIFLTPTIYFIIFLTGIVGNGLVILVM 63
QY 66 GYOKKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIIYTVNLVSS 125
Db 64 GYOKKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIIYTVNLVSS 123
QY 126 VLILAFISLDRLYLAIVHATNSORPRKLLAEKAVYGVWIPALLLTIPDFIFADVSGQDIS 185
Db 124 VLILAFISLDRLYLAIVHATNSORPRKLLAEKAVYGVWIPALLLTIPDFIFADVSGQDIS 178
QY 186 QGDDRYICDRLYPDSLMVMVVFQCHIMVGLIIPGIVILSCYCIISKLSSHKGHOKRKAL 245
Db 179 EADDRYICDRFYENDLWVVVFQCHIMVGLIIPGIVILSCYCIISKLSSHKGHOKRKAL 238
QY 246 KTTVILILAFFACWLPYVIGISIDSPTILGVIKQGDFFSIVHKWISITEALAFFHCCLN 305
Db 239 KTTVILILAFFACWLPYVIGISIDSPTILGVIKQGDFFSIVHKWISITEALAFFHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALNSMRGSSSLKILSKRGHSHSVSTESSESSPHSS 359
Db 299 PILYAFILGAKFKTSQAHALTSVRGSSSLKILSKRGHSHSVSTESSESSPHSS 352

Search completed: May 17, 2004, 22:38:46
Job time : 549 secs

QY 6 VSIYTSNDYSEVSGSDYDSNKEPCFRDENVHFNRIFLTPTIYFIIFLTGIVGNGLVILVM 65
Db 4 ISIYTSNDYSEVSGSDYDSNKEPCFRDENVHFNRIFLTPTIYFIIFLTGIVGNGLVILVM 63
QY 66 GYOKKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIIYTVNLVSS 125
Db 64 GYOKKLRSMTDKYRLHLSVADLLFVITLFPWADAMADWYFGKFLCKAVHIIYTVNLVSS 123
QY 126 VLILAFISLDRLYLAIVHATNSORPRKLLAEKAVYGVWIPALLLTIPDFIFADVSGQDIS 185
Db 124 VLILAFISLDRLYLAIVHATNSORPRKLLAEKAVYGVWIPALLLTIPDFIFADVSGQDIS 178
QY 186 QGDDRYICDRLYPDSLMVMVVFQCHIMVGLIIPGIVILSCYCIISKLSSHKGHOKRKAL 245
Db 179 EADDRYICDRFYENDLWVVVFQCHIMVGLIIPGIVILSCYCIISKLSSHKGHOKRKAL 238
QY 246 KTTVILILAFFACWLPYVIGISIDSPTILGVIKQGDFFSIVHKWISITEALAFFHCCLN 305
Db 239 KTTVILILAFFACWLPYVIGISIDSPTILGVIKQGDFFSIVHKWISITEALAFFHCCLN 298
QY 306 PILYAFILGAKFKSSAQHALNSMRGSSSLKILSKRGHSHSVSTESSESSPHSS 359
Db 299 PILYAFILGAKFKTSQAHALTSVRGSSSLKILSKRGHSHSVSTESSESSPHSS 352

RESULT 15
US-10-239-423-79
; Sequence 79, Application US/10239423
; Publication No. US20030186889A1
; GENERAL INFORMATION:
; APPLICANT: FORSSMANN, Wolf-Georg; FORSSMANN, Ulf; ADERMAN, Knut;
; TITLE OF INVENTION: Diagnostic Agent and Medicament for Examining the
; TITLE OF INVENTION: Cell Surface Proteome of Tumor and Inflammation Cells and
; TITLE OF INVENTION: for Treating Tumor Diseases and Inflammatory Diseases,
; TITLE OF INVENTION: Preferably with the Aid of Specific Chemokine
; TITLE OF INVENTION: Receptor Analysis and Chemokine Receptor/Ligand Interaction
; FILE REFERENCE: 022217us
; CURRENT APPLICATION NUMBER: US/10/239,423
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: DE10016013.1
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Amino Acid Sequence for the Generation of Antibodies
US-10-239-423-79

Query Match 89.9%; Score 1678.5; DB 14; Length 352;
Best Local Similarity 90.1%; Pred. No. 1.e-145;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2004, 21:33:47 ; Search time 44 Seconds
(without alignments)
784.835 Million cell updates/sec

Title: US-09-367-052-2

Perfect score: 1867
Sequence: 1 MEISVSIYSDNYSEEVGS.....KRGHSSVTESESSSFHSS 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_78: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1679.5	90.0	352	G00048	fusin (LESTRA) - c
2	1678.5	89.9	352	A45747	neuropeptide Y/pep
3	1631	87.4	353	S28787	neuropeptide Y/pep
4	585.5	31.4	367	J80349	interferon-inducib
5	570.5	30.6	360	A57160	chemokine (C-C) re
6	551.5	29.5	355	JQ1231	interleukin-8 rece
7	550.5	29.5	378	A55735	G protein-coupled
8	547.5	29.3	356	S42096	interleukin-8 rece
9	544.5	29.2	358	A53752	interleukin-8 rece
10	544.5	29.2	360	A53611	interleukin-8 rece
11	537.5	28.8	378	B55735	lymphocyte-specifi
12	535	28.7	369	JC5068	G protein-coupled
13	533	28.5	355	JC4304	orphan G protein-c
14	530	28.4	360	JC4587	chemokine (C-C) re
15	526.5	28.2	327	S95162	MDCR15 protein - h
16	523.5	28.0	359	A48921	interleukin-8 rece
17	522.5	28.0	350	A39445	interleukin-8 rece
18	522.5	28.0	372	S26667	G protein-coupled
19	518.5	27.8	378	A45680	G protein-coupled
20	517.5	27.7	374	S42628	G protein-coupled
21	508.5	27.2	374	S32785	G protein-coupled
22	507	27.2	383	S55594	G protein-coupled
23	501	26.8	354	S58186	probable G protein
24	496	26.6	375	J18450	chemokine (C-C) re
25	494	26.5	355	JC5067	G protein-coupled
26	484	25.9	350	JN0621	G protein-coupled
27	483.5	25.9	352	A43113	chemokine (C-C) re
28	479.5	25.7	360	JC2443	chemokine (C-C) re
29	476.5	25.5	355	G02436	chemokine (C-C) re

30	473.5	25.4	359	2	S15403	angiotensin II rec
31	472.5	25.3	355	2	A45177	chemokine (C-C) re
32	463.5	24.8	359	2	JC1104	angiotensin II rec
33	463.5	24.8	362	2	JN0694	angiotensin II rec
34	460	24.6	362	2	A30341	G protein-coupled
35	459.5	24.6	359	2	S44425	angiotensin II rec
36	459.5	24.6	359	2	A42656	angiotensin II rec
37	459.5	24.6	359	2	JH0621	angiotensin II rec
38	459.5	24.6	359	2	JC1134	angiotensin II rec
39	457.5	24.5	359	2	A48857	angiotensin II rec
40	453.5	24.3	359	2	JQ1516	angiotensin II rec
41	453	24.3	354	2	A23669	interleukin-8 rece
42	448	24.0	362	2	A39714	G protein-coupled
43	447.5	24.0	359	2	I49341	MIP-1 alpha recept
44	445.5	23.9	359	2	I39418	angiotensin II rec
45	442.5	23.7	359	2	JC1194	angiotensin II rec

ALIGNMENTS

RESULT 1

G00048
fusin (LESTRA) - crab-eating macaque
C/Species: Macaca fascicularis (crab-eating macaque)
C/Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 26-Aug-1999
C/Accession: G00048
R/Tatsumi, M.
submitted to GenBank, July 1996
A/Reference number: H00048
A/Accession: G00048
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-352 <TAT>
A/Cross-references: GB:D86579; MID:G1468948; PID:G1468949
C/Superfamily: vertebrate rhodopsin

Query Match 90.0%; Score 1679.5; DB 2; Length 352;
Best Local Similarity 90.1%; Pred. No. 5.3e-140;
Matches 319; Conservative 17; Mismatches 13; Indels 5; Gaps 1;

QY	6	VSIIYSDNYSEVSGDYDSNKPCPRDENVFNRIPLTYIFLPTIYFIIVGNGVILVM	65
DB	4	ISIIYSDNYTEMSGDYDSIKPCPCRENAHFNRIPLTYISIIIFLPTIYFIIVGNGVILVM	63
QY	66	GYQKLRMTDKYRLHLSVADLLFVITLFWAVDAMADWFGKFLCKAVHIYTVNLVSS	125
DB	64	GYQKLRMTDKYRLHLSVADLLYVITLFWAVDANWYFGNFKCKAVHIYTVNLVSS	123
QY	126	VLILAFISLDRYLATVHATNSQRPKLLAEKAVYVGVWIPALLTIPDFPADVSGQDIS	185
DB	124	VLILAFISLDRYLATVHATNSQRPKLLAEKAVYVGVWIPALLTIPDFIFASV----	178
QY	186	QGDDRYICDRLYPDSLWMVVFQFHIMVGLILPGVILSCYCIISKLSHSGHOKRKAL	245
DB	179	EADRYICDRFPNDLWVVFQFHIMVGLILPGVILSCYCIISKLSHSGHOKRKAL	238
QY	246	KTTVILILAFACWLPYVIGISIDSFILLGVKQCFDESIVHKWISITEALAFPHCLN	305
DB	239	KTTVILILAFACWLPYVIGISIDSFILLEIKQCFEPTVHKWISITEALGFPHCLN	298
QY	306	PLYIAFLGAKPKSSCHALNSMRGSSIKILSKGRGSHSSVTESESSSFHSS	359
DB	299	PLYIAFLGAKFKTSQAHALTSVSRGSSIKILSKGRGSHSSVTESESSSFHSS	352

RESULT 2

A45747
neuropeptide Y/peptide YY receptor Y3 - human
N/Alternate names: fusin; HM89; leukocyte-derived seven-transmembrane receptor LESTR; re
C/Species: Homo sapiens (man)
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-Nov-1999
C/Accession: A45747; A53103; I53006; I59444; I69203; S32761


```
QY 17 EVGSDYD-SNKPCFEDNVHNFRIPLTIPIITGIVGNGLVILVNGYQKLSMT 75
Db 28 DYGENESDSDPCPDQFSLNDRTEPLPALYSLLFLGLGNGAVALLSORTALST 87
QY 76 DKYRLHLSVADLLFVITLFPWADAMADYFGLFKCAVHIIYTNVLSVLLAFISLD 135
Db 88 DTFLLHLAVADLLVLTPLWADAAVQWVFGPLCKVAGALNFNIFAGAFLLACISFD 147
QY 136 RYLAIIVHATNSQR--PRKLLAEKAVVGVWIPALLITIPDFIADVSQGDSDRYI- 192
Db 148 RYLSIVHATQYRRDPRVRVAVLCIV--VWGLCLLFPALPDFIY-----LSANYDORLN 198
QY 193 ---CDRLYPDSLMMVVFQFHIMVGLILPGIVILSVLCYCIISKLKSGHQRKALKTTV 249
Db 199 ATHCOYNFP-QVGRGTAALRVQLVAGFLLPVLVAYCYAHILAVLLVSRGQRRFRANLVV 257
QY 250 ILILAPFACWLPYVVGISIDSIFILLGVKOGDPFESIVHKWISITEALAFHCCLNPIIY 309
Db 258 VVVAAPAVCVTHLVVLDVLDVGLARNCGRSHVDVAKSVTSGMGYMHCLNPIIY 317
QY 310 AFLGAKFKSSAQHALNSMRGSSLSKILKSGHGHSSVSTESSSF 356
Db 318 AFVGKREQWMLFTRLGSD-----QRGQRPQSSRRRESSW 356

RESULT 5
A57160
chemokine (C-C) receptor 4 - human
N:Alternate names: C-C CKR-4
C:Species: Homo sapiens (man)
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: A57160
R:Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogwerf, A.J.; Proudfoot, A.E.I.; W
J. Biol. Chem. 270, 19495-19500, 1995
A:Title: Molecular cloning and functional expression of a novel CC chemokine receptor cD
A:Reference number: A57160; MUID:95370289; PMID:7642634
A:Accession: A57160
A:Status: Preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-360 <POW>
A:Cross-references: GB:X85740; NID:G1370103; PIDN:CAA59743.1; PID:g971452
A:Note: source clone X5-5
C:Genetics:
A:Gene: GDB:CMKBR4
A:Cross-references: GDB:677463
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:40-65/Domain: transmembrane #status predicted <TM1>
F:76-97/Domain: transmembrane #status predicted <TM2>
F:112-133/Domain: transmembrane #status predicted <TM3>
F:151-175/Domain: transmembrane #status predicted <TM4>
F:208-226/Domain: transmembrane #status predicted <TM5>
F:243-264/Domain: transmembrane #status predicted <TM6>
F:291-308/Domain: transmembrane #status predicted <TM7>
F:29-276, 110-187/Disulfide bonds: #status predicted
F:72,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
F:143/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:183,194/Binding site: carboxylate (Asn) (covalent) #status predicted
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 30.6%; Score 570.5; DB 2; Length 360;
Best Local Similarity 37.4%; Pred. No. 1.7e-42;
Matches 120; Conservative 63; Mismatches 125; Indels 13; Gaps 6;

QY 1 MEPISVIYTSNDYSEVSGD--YDSNKEPCFEDNVHNFRIPLTIPIITGIVGN 58
Db 1 MNPTDIADTTLD---ESIYNNYLYESI PKPKTEKIGIKAFGELFPLPLYSLVFVGLGN 57
QY 59 GLVILVNGYQKLSMTDKYRLHLSVADLLFVITLFPWADAMADYFGLFKCAVHIIY 118
Db 58 SVVVLVFKYKRKLSMTDVLNLAISDLLFVFLSPWGYAADQWVFGGLCKWISWY 117
```

```
QY 119 TNVLSVLLIAPISLDRIYLAIVHATNSQRPRKLLAEKAVVGVWIPALLITIPDFIPAD 178
Db 118 LVGFYSGIFVYMLASIDRIYLAIVHAFVSRLARTLTITGYITSLATWSVAVFASLPQFLFST 177
QY 179 VSQGDISQGDRIYCDRLYP--DSLMMVVFQFHIMVGLILPGIVILSVLCYCIISKLKSHS 236
Db 178 C-----YTERNHY-CKTKYSLNSTWKLSSLEINILGLVPLGIMLFYCSWIRTLQHC 232
QY 237 KHKQRKALKTTVILILAPFACWLPYVVGISIDSIFILLGVKOGDPFESIVHKWISITEA 296
Db 233 KNEKKNKAVKIMFAVVVLFGFWTPYINVLFLFLETLVEVL-ODCTERYLDYIAQATET 291
QY 297 LAFFHCCLNPIIYAFGLGAKPK 317
Db 292 LAFVHCCLNPIIYFGLGEKFR 312

RESULT 6
JQ1231
interleukin-8 receptor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999
C:Accession: JQ1231; A46483
R:Beckmann, M.P.; Murger, W.E.; Kozlosky, C.; VandenBos, T.; Price, V.; Lyman, S.; Gerar
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A:Title: Molecular characterization of the interleukin-8 receptor.
A:Reference number: JQ1231; MUID:91378994; PMID:1898400
A:Accession: JQ1231
A:Molecule type: DNA
A:Residues: 1-355 <BEC>
A:Cross-references: GB:M74240; NID:g165439; PIDN:AAA31375.1; PID:g165439
R:Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A:Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.
A:Reference number: A46483; MUID:92148149; PMID:1737938
A:Accession: A46483
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-355 <LEE>
A:Cross-references: GB:M82873; NID:g165440; PIDN:AAA31376.1; PID:g165441
A:Experimental source: neutrophils
A:Note: sequence extracted from NCBI backbone (NCBIN:81526, NCBI:81530)
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 29.5%; Score 551.5; DB 2; Length 355;
Best Local Similarity 35.0%; Pred. No. 7.7e-41;
Matches 121; Conservative 72; Mismatches 142; Indels 11; Gaps 6;

QY 12 DNYSEVSGDYDSNKEPCFEDNVHNFRIPLTIPIITGIVGNGLVILVNGYQKKL 71
Db 17 DEFANATGMPPEVKDYSPLVVTQT-LNKYVYVVIYALVFLSLGSLVLMVLVYGRSN 75
QY 72 RSMTDKYRLHLSVADLLFVITLFPWADAMADYFGLFKCAVHIIYTNVLSVLLIAP 131
Db 76 RSVTDVILNLAADLLFALTPTWASKEGWTGTPCLKVSLVKEVNFYSGLILLAC 135
QY 132 ISLDRIYLAIVHATNSQRPRKLLAEKAVVGVWIPALLITIPDFIADVSQGDISQGDRIY 191
Db 136 ISVDRIYLAIVHATRTLTQKRLV-KFICLGIWALSLLSLPFLFRQV---PSPNNSSP 190
QY 192 IC--DRLYPDSLMMVVFQFHIMVGLILPGIVILSVLCYCIISKLKSHSGHQRKALKTTV 249
Db 191 VCYEDLGHTAKRWLVRLPHTFTGFLPLVLMVFCYGTTLTTFQAHMGOKHRMVRIF 250
QY 250 ILILAPFACWLPYVVGISIDSIFILLGVKOGDPFESIVHKWISITEALAFHCCLNPIIY 309
Db 251 AVVLIFFLCWLPYNVLLADLTMTHTVYIQTCCQRNDIDRALDATEILGFLHSCINPIIY 310
QY 310 AFLGAKFKSSAQHALNSMRG--SSKLKSKGRGSHSVSTESSS 354
Db 311 AFIQGNFNGFLKML--AAGRLISKEFLTRHRTVTSYSSSTNTPSN 354
```

```
RESULT 7
A:Accession: A55735
C:Species: Mus musculus (house mouse)
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Nov-1999
C:Accession: A55735
R:Schweickart, V.L.; Raport, C.J.; Godiska, R.; Byers, M.G.; Eddy Jr., R.L.; Shows, T.B.
Genomics 23, 643-650, 1994
A:Title: Cloning of human and mouse EB1, a lymphoid-specific G-protein-coupled receptor
A:Reference number: A55735; MUID:195154835; PMID:7851893
A:Accession: A55735
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <SCH>
A:Cross-references: GB:131580; NID:9468340; PIDN:AAA74232.1; PID:9468341
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor

Query Match 29.5%; Score 550.5; DB 2; Length 378;
Best Local Similarity 33.2%; Pred. No. 1e-40;
Matches 117; Conservative 78; Mismatches 150; Indels 7; Gaps 3;

Qy 11 SDNYSEVGSQDYDNKPCFRDENVHFNRIPLTIYFIPLTGIVGNGLVILVNGYQK 70
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 29 TDDYIGENTVDYTLYESVCFKDVNRKAFPLMYSVCFVGLLGNGLVILTIYVPR 88
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 71 LRSMTDKYRLHLSVADLLFVITLPFWADAMADWYFGKFLCKAVHIIYTVNLVSVLILA 130
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 89 LKTMIDTYLNLAVADILFLLPLFWAYSEAKSWFGVLCGIGFIYKLSFFSGMLLL 148
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 131 FLSLDRYLAIVHAT--NSQRPRLAEKAVYGVWIPALLTIPTDFPADVSQGDISQGD 188
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 149 CISIDRYVAIVQAVSRHRRARVLLISLSCVGIWMLALFLSIPELLYSGLQK--NSGE 205
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 189 DRYICDRLYPSDLWVWVFOHIMVGLILPGIIVLSVCIIISKLSSHKGKQKALKTT 248
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 206 DTLRSLVSAQVEALITIQVQVFGVFLVLMAMGFCYLIIRTLQARNFERNKAIKY 265
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 249 VILILAFFACWLPYVVGISIDSFILLGVIKQCDPESIVHKWISITALEAFHCCNLPIL 308
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 266 IAWVVFVIFQLPYNGVLAQTVANFNITNSCETSQKLNIAVDYTVSLASVRCVNPFL 325
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 309 YAPLGNKPKVSSQAQHALNSMRSS--LKILSKGKGGHSSVSTESSESSPFS 358
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 326 YAFIGVKFRSDFLKFUDKGLCLSQERLRHWSRCHRVNASVSMEATTTTTS 377
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 8
S42096
interleukin-8 receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C:Accession: S42096
R:Gobl, A.E.; Wang, S.; Zhou, Y.; Oeborg, K.
submitted to the EMBL Data Library, February 1994
A:Description: Molecular cloning of the rat IL8 receptor.
A:Reference number: S42096
A:Accession: S42096
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-356 <GOB>
A:Cross-references: EMBL:X77797
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 29.3%; Score 547.5; DB 2; Length 356;
Best Local Similarity 34.8%; Pred. No. 1.7e-40;
Matches 129; Conservative 67; Mismatches 124; Indels 51; Gaps 11;

Qy 12 DNYS-EEVCSGDYD-----SNKEPCFRDENVHFNRIPLTIYFIPLTGIVGN 58
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 7 DNFSLEDFSGSDYDYNFSSDPFTLSDAAPC-PSANULDINRYAVVVIYVILTLVLGN 65
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

```
Qy 59 GLVILVNGYQKLRSMTDKYRLHLSVADLLFVITLPFWADAMADWYFGKFLCKAVHIIY 118
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 119 TVNLVSVLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPTDFPAD 178
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 126 EIFFYSSVLLACISNDRYLAIVHATSTLIQKHLV-KFVCITTMWFLSVLSLPIFIL-- 182
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 179 VSGDISQGDRIYCDRLYPD-----SLWVWVFOHIMVGLILPGIIVLSVCIIISKL 233
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 183 --RTTVKANPSTVVC---YENIGNNTSKRWVLRILPQTYGFLFLPLLMFCYGFTRLTL 237
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 234 SHSKGHQKQKALKTTVILILAFACWLPYVVGISIDSFILLGVIKQCDPESIVHKWISI 293
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 238 FKAMQKQKRAMVFAVVLVFLCWLFPYVILVFTDTLMRTKLKETCERQNEINK--A 294
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 294 TEALAFHCCNLPILYAFIAGKPKSSAQHALNSMRGSSLSKILSK-----GKGGH 344
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 295 SEILGFLHSLCLNPIIYAFIQKFR----HGL-----LKIMANYGLSVKEFLAKEGRP 342
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 345 SSVSTESSESS 355
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 9
A53752
interleukin-8 receptor (clone 5Bla) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: A53752
R:Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro,
J. Biol. Chem. 269, 12391-12394, 1994
A:Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A:Reference number: A53752; MUID:94230294; PMID:8175642
A:Accession: A53752
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-358 <PRA>
A:Cross-references: GB:I24445; NID:9437661; PIDN:AAA31378.1; PID:9437662
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 29.2%; Score 544.5; DB 2; Length 358;
Best Local Similarity 34.7%; Pred. No. 3.2e-40;
Matches 128; Conservative 70; Mismatches 132; Indels 39; Gaps 9;

Qy 9 YTSQDYSEVGSQDYDS-----NKEPCFRDENVHFNRIPLTIYFIPLTGIV 56
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 4 FTWENSYEDFFGDFSNYSVSTDLPTLLDSAPC-RSESLTNSYVVLITYILVFLLSLL 62
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 57 GNGLVILVNGYQKLRSMTDKYRLHLSVADLLFVITLPFWADAMADWYFGKFLCKAVHI 116
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 63 GNSLWMLVILYSRSTCSTVDVYLLNLAIADLLPATTLP-MAASKVHGWTGTPCLKVWSL 122
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 117 IYTVNLVSVLILAFISLDRLYLAIVHATNSQRPRLAEKAVYGVWIPALLTIPTDFIF 176
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 123 VKEVNFYSGILLACISVDRYLAIVHATRTMIQKHLV-KFICLSMWGVSLLISLILF 181
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 177 ADVSQGDISQDDRIYCDRLYPD-----SLWVWVFOHIMVGLILPGIIVLSVCIIIS 231
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 182 RNA----IPPNSSPVC---YEDMGNSTAKRWVLRILPQTFGFIPLVLMVLCYVFTLR 234
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 232 KLSHKGKQKQKALKTTVILILAFACWLPYVVGISIDSFILLGVIKQCDPESIVHKWI 291
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 235 TLFQAHMGQKQKRAMVFAVVLFLCWLFPYVILVLTDTLMRTHVQETCERNIDRAL 294
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 292 SITEALAFFHCCNLPILYAFIAGKAF-----KSSAQHALNSMRGSSLSKILSKGKGGHSS 346
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 295 DATEILGFLHSLCLNPIIYAFIQKFRYGLLKIILAHGL-----ISKEFLAKESR--PSF 346
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 347 VSTESSESS 355
Db :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```


JC5068
G protein-coupled receptor CKR-L3 - human
C/Species: Homo sapiens (man)
C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 21-Jul-2000
C/Accession: JC5068
R:Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
Biochem. Biophys. Res. Commun. 227, 846-853, 1996
A/Title: Molecular cloning and RNA expression of two new human chemokine receptor-like genes
A/Reference number: JC5067; MUID:97040707; PMID:8886020
A/Accession: JC5068
A/Molecule type: DNA
A/Residues: 1-369 <ZAB>
A/Cross-references: ENBL:279784; NID:G1668737; PIDN:CA02144.1; PID:G1668738
C/Comment: This protein belongs to the family of alpha chemokine receptors.
C/Genetics:
A/Genes: GDB:CMK66; SREL22; GPR29; CCR6; CKR-L3; GPR-CY4
A/Cross-references: GDB:5370639; OMIM:601835
A/Map position: 6q27-6q27
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein
F:42-68/Domain: transmembrane #status predicted <TM1>
F:79-99/Domain: transmembrane #status predicted <TM2>
F:115-136/Domain: transmembrane #status predicted <TM3>
F:160-180/Domain: transmembrane #status predicted <TM4>
F:212-233/Domain: transmembrane #status predicted <TM5>
F:250-271/Domain: transmembrane #status predicted <TM6>
F:292-315/Domain: transmembrane #status predicted <TM7>
Query Match 28.7%; Score 535; DB 2; Length 369;
Best Local Similarity 33.9%; Pred. No. 2.3e-39;
Matches 124; Conservative 72; Mismatches 136; Indels 34; Gaps 9;
QY 10 TSDNYSEEVGSDY--DSNKPCPRDENVHFNRIPLPIYFIILTGIVGNGLVILVNGY 67
DB 9 SSEDYFVSNTSYSDVSEMLLCSLQEVRFSLFVPLAYSLLICVFGLLGNILVITFAF 68
QY 68 QKLSMTDKYRLHLSVADLLFVITLPFVAVD-AMADWVFGKFLCKAVHIITVNLVSSV 126
DB 69 YKRSMTDVIYLNWAIADILFVITLPFVAVSHATGAWFVSNATKLDKGIYAFNFCGM 128
QY 127 LILATISLDRIYLAIVHATNS--QRPRKLLAEKAVYGVWIPALLTIPDFIFADVSGDI 184
DB 129 LLITCISMDRYAIYQATKSPRLSRTLPKSKICLVVWGLSVIISSTFVF--NOKYN 185
QY 185 SGDDRYICDRLY-----PDSLWVVFQHIWVGLILPGVILSCYCIILSKLSHGKHQ 240
DB 186 TQSGD--VCEPKYQTVSEPIRKLMLGLLELFGFFILPMFMYCTFIVTKLVQAQNSK 243
QY 241 KKKALKTTVILIAFFACWLPYVVGISIDSPFLLGVKQGDFFSIVHKWISITEALAF 300
DB 244 RKAIRVIAVVLVLAQIPHWVLLVTA-ANLGNWNSCQSEKLGITKTVTEVLAF 302
QY 301 HCLNPLIYAFLGAKFKSAQHALNSMRGSSLKILS-----KGRGSHSVSTSES 353
DB 303 HCLNPLVYAFITGQKF-----RNVFLKILKDLWCVRKYKSGSFSCAGRYSEN 350
QY 354 SPSFAS 359
DB 351 ISRQTS 356
RESULT 13
JC4304
Orphan G protein-coupled receptor - human
N/Alternate names: V28 protein
C/Species: Homo sapiens (man)
C/Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 19-May-2000
C/Accession: JC4304
R:Raport, C.J.; Schweickart, V.L.; Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
Gene 163, 295-299, 1995
A/Title: The orphan G-protein-coupled receptor-encoding gene V28 is closely related to G
A/Reference number: JC4304; MUID:96011651; PMID:7590284
A/Accession: JC4304

A/Molecule type: mRNA
A/Residues: 1-355 <RAP>
A/Cross-references: GB:U20350; NID:G665580; PIDN:AAA91783.1; PID:G665581
A/Experimental source: peripheral blood mononuclear cell
C/Comment: This protein is a cell-surface receptor which recognizes extracellular signals;
C/Comment: This protein is a key regulator of many immune and homeostatic responses, and
C/Genetics:
A/Genes: V28
A/Map position: 3pter-p21
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; lymphokine; transmembrane protein
F:35-57/Domain: transmembrane #status predicted <TM1>
F:66-88/Domain: transmembrane #status predicted <TM2>
F:104-125/Domain: transmembrane #status predicted <TM3>
F:146-165/Domain: transmembrane #status predicted <TM4>
F:197-217/Domain: transmembrane #status predicted <TM5>
F:230-254/Domain: transmembrane #status predicted <TM6>
F:275-296/Domain: transmembrane #status predicted <TM7>
Query Match 28.5%; Score 533; DB 2; Length 355;
Best Local Similarity 33.1%; Pred. No. 3.2e-39;
Matches 120; Conservative 65; Mismatches 146; Indels 32; Gaps 7;
QY 12 DNYSEEVGSG-DYDSNKPCPRDENVHFNRIPLPIYFIILTGIVGNGLVILVNGYQKK 70
DB 2 DOPFESVTENFEYDDLAEACYIGDIVFGTVFSLSYFVIFAIGLVGNLLVVFALTNSKK 61
QY 71 LRSMTDKYRLHLSVADLLFVITLPFVAVDAMADWVFGKFLCKAVHIITVNLVSSVILA 130
DB 62 PKSVTDIYLNALSLDLFVATLPFWTHYLNEKGLHNAMCKFTTAFPIGFSGSIFIT 121
QY 131 FISLDRIYLAIVHATNSQRPRKLLAEKAVYGVWIPALLTIPDFIFADVSGQDISQDDR 190
DB 122 VISIDRYLAIVLAANSMMNRITVQHGVTISLGVAAAILVAAPQFMPTKOKENE----- 174
QY 191 YICDRLYPD---SLWVVFQHIWVGLILPGVILSCYCIILSKLSHGKHQKALKKT 247
DB 175 --CLGDPYEVQLQEIWVLRVNETNFGFLPLILINSYCYFRIIQTLSCKNHKAKAIKL 232
QY 248 TVTILIAFACWLPYVVGISIDSPFLLGVKQGDFFSIVHKWISITEALAFHCCLNFI 307
DB 233 ILVAVVFFLFTPTVNMIFLETKLIDYDF-PSCDMRKLRLALSIVTETVAFSHCCLNFI 291
QY 308 LYAFLGAKFKSAQHALNSMRGSSLKIL-----SKGRGSHSVSTSESSESSF 356
DB 292 IYAFAGEKPRVLYHLY-----GKCLAVLCGRSVHVDFFSSSSQSRHGSVL--SSNFTY 344
QY 357 HSS 359
DB 345 HTS 347
RESULT 14
JC4587
Chemokine (C-C) receptor 4 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 08-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jun-2000
C/Accession: JC4587
R:Hoogwerf, A.J.; Black, D.; Proudfoot, A.E.I.; Wells, T.N.C.; Power, C.A.
Biochem. Biophys. Res. Commun. 218, 337-343, 1996
A/Title: Molecular cloning of murine CC CKR-4 and high affinity binding of chemokines to
A/Reference number: JC4587; MUID:96136324; PMID:8573157
A/Accession: JC4587
A/Molecule type: mRNA
A/Residues: 1-360 <HOO>
A/Cross-references: EMBL:X90862; NID:G1167851; PIDN:CAA62372.1; PID:G1167852
A/Experimental source: thymus
C/Genetics:
A/Genes: CC ckr-4
C/Superfamily: vertebrate rhodopsin
C/Keywords: Glycoprotein; phosphoprotein; receptor; thymus
F:2,183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:72,202,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predi

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 17, 2004, 16:16:00 ; Search time 27 Seconds
(without alignments)
692.340 Million cell updates/sec

Title: US-09-367-052-2

Perfect score: 1867
Sequence: 1 MEFISVSIYSDNYSEVGS.....KRGCHSSVTESESSFFHSS 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1867	100.0	359	1 CCR4_MOUSE	P70658 m C-X-c che
2	1744.5	93.4	349	1 CCR4_RAT	O08565 rattus norv
3	1686.5	90.3	352	1 CCR4_PAPAN	P56491 papio anubi
4	1679.5	90.0	352	1 CCR4_MACFA	Q28474 macaca fasc
5	1678.5	89.9	352	1 CCR4_HUMAN	P30991 homo sapien
6	1676.5	89.8	352	1 CCR4_MACMU	P79394 macaca mula
7	1668.5	89.4	352	1 CCR4_CERTO	O62747 cercopithec
8	1643	88.0	353	1 CCR4_FELCA	P56498 felis silve
9	1631	87.4	353	1 CCR4_BOVIN	P25930 bos taurus
10	885.5	47.4	192	1 CCR4_SHEEP	Q28553 ovis aries
11	584	31.3	368	1 CCR3_HUMAN	P49682 homo sapien
12	581.5	31.1	367	1 CCR3_MOUSE	O88410 mus musculu
13	570.5	30.6	360	1 CCR4_HUMAN	P51679 homo sapien
14	562	30.1	359	1 IL8B_RAT	P35407 rattus norv
15	551.5	29.5	355	1 IL8B_RABIT	P21109 oryctolagus
16	550.5	29.5	378	1 CCR7_MOUSE	P47774 mus musculu
17	548.5	29.4	356	1 IL8B_CANFA	O97571 canis famil
18	544.5	29.2	358	1 IL8B_RABIT	P35344 oryctolagus
19	544.5	29.2	360	1 IL8B_HUMAN	P25025 homo sapien
20	542.5	29.1	353	1 IL8B_PANTR	Q28807 pan troglod
21	541.5	29.0	353	1 IL8B_MACMU	Q28519 macaca mula
22	541	29.0	353	1 IL8B_GORGO	Q28422 gorilla gor
23	537.5	28.8	378	1 CCR7_HUMAN	P32248 homo sapien
24	536	28.7	369	1 CCR9_MOUSE	O98077 mus musculu
25	535	28.7	374	1 CCR6_HUMAN	P51684 homo sapien
26	534	28.6	357	1 CCR9_HUMAN	P51686 homo sapien
27	533	28.5	355	1 CXX1_HUMAN	P49238 homo sapien
28	532	28.5	367	1 CCR6_MOUSE	O54689 mus musculu
29	530	28.4	360	1 CCR4_MOUSE	P51680 mus musculu
30	528	28.3	378	1 CCR6_MOUSE	O08707 mus musculu
31	525.5	28.1	350	1 IL8A_GORGO	P55919 gorilla gor
32	524.5	28.1	350	1 IL8A_PANTR	P55920 pan troglod
33	523.5	28.0	359	1 IL8B_MOUSE	P35343 mus musculu

ALIGNMENTS

RESULT 1

ID	CCR4_MOUSE	STANDARD;	PRT;	359 AA.
AC	P70658; O09059; O09062; P70233; P70346;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (Stromal cell-			
DE	derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived			
DE	seven transmembrane domain receptor) (LESTR) (Pre-B-cell-derived			
DE	chemokine receptor) (PB-CR).			
GN	CXCR4 OR LESTR OR CMKAR4 OR SDF1R.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
[1]				
RP	SEQUENCE FROM N.A. (ISOFORM CXCR4-B).			
RC	STRAIN=C57BL/6J, and 129/SV; TISSUE=Peritoneal exudate;			
RX	MEDLINE=97113334; PubMed=8955194;			
RA	Heesen M., Berman M.A., Benson J.D., Gexard C., Dorf M.E.;			
RT	"Cloning of the mouse fusin gene, homologue to a human HIV-1			
RT	co-factor.";			
RL	J. Immunol. 157:5455-5460(1996).			
[2]				
RN	SEQUENCE FROM N.A. (ISOFORM CXCR4-B).			
RC	TISSUE=Pre-B cell;			
RX	MEDLINE=97121456; PubMed=8962122;			
RA	Nagasawa T., Nakajima T., Tachibana K., Iizasa H., Bleul C.C.,			
RA	Yoshie O., Macushima K., Yoshida N., Springer T.A., Kishimoto T.;			
RT	"Molecular cloning and characterization of a murine pre-B-cell			
RT	growth-stimulating factor/stromal cell-derived factor 1 receptor, a			
RT	murine homolog of the human immunodeficiency virus 1 entry coreceptor			
RT	fusin.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:14726-14729(1996).			
[3]				
RN	SEQUENCE FROM N.A. (ISOFORM CXCR4-B).			
RC	STRAIN=129/SV; TISSUE=Thymus;			
RA	Schubel A., Burgstahler R., Lipp M.;			
RT	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.			
[4]				
RN	SEQUENCE FROM N.A. (ISOFORMS CXCR4-A AND CXCR4-B).			
RC	STRAIN=C57BL/6J x CBA; TISSUE=Thymus;			
RX	MEDLINE=97439495; PubMed=9295051;			
RA	Moepes B., Frodl R., Rodewald H.-R., Baggiolini M., Gierschik P.;			
RT	"Two murine homologues of the human chemokine receptor CXCR4 mediating			
RT	stromal cell-derived factor 1alpha activation of G12 are			
RT	differentially expressed in vivo.";			
RL	Eur. J. Immunol. 27:2102-2112(1997).			
[5]				
RN	SEQUENCE FROM N.A. (ISOFORMS CXCR4-A AND CXCR4-B).			
RC	MEDLINE=97256574; PubMed=9103415;			
RA	Heesen M., Berman M.A., Hoepken U.E., Gerard N.P., Dorf M.E.;			
RT	"Alternate splicing of mouse fusin/CXC chemokine receptor-4: stromal			
RT	cell-derived factor-1alpha is a ligand for both CXC chemokine			
RT	receptor-4 isoforms.";			

34	522.5	28.0	372	1 CCR5_HUMAN	P33302 homo sapien
35	520.5	27.9	350	1 IL8A_HUMAN	P25024 homo sapien
36	517.5	27.7	374	1 CCR5_MOUSE	O04683 mus musculu
37	512.5	27.5	360	1 IL8B_BOVIN	O28003 bos taurus
38	511.5	27.4	353	1 CCR8_MOUSE	P56484 mus musculu
39	510	27.3	342	1 CCR6_HUMAN	O09574 homo sapien
40	510	27.3	342	1 CCR6_PANTR	Q9CV16 pan troglod
41	508.5	27.2	374	1 CCR5_RAT	P34997 rattus norv
42	508	27.2	382	1 CKD6_RAT	O09027 rattus norv
43	505.5	27.1	342	1 CCR6_CERAE	O18983 cercopitheci
44	504	27.0	342	1 CCR6_MACNE	O19024 macaca neme
45	501	26.8	354	1 CXX1_RAT	P35411 rattus norv

RL J. Immunol. 158:3561-3564 (1997).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RX STRAIN=C57BL/6; TISSUE=Thymus;
RA Suzuki G., Nakata Y., Uzawa A., Shirasawa T., Saito T., Mita K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM CXCR4-B).
RX STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [8]
RP ALTERNATIVE SPLICING.
RX MEDLINE=98295994; PubMed=9634237;
RX MEDLINE=99095114; PubMed=9879064;
RP Prodi R., Gierschik P., Moepf B.;
RT "Genomic organization and expression of the CXCR4 gene in mouse and
RT man: absence of a splice variant corresponding to mouse CXCR4-B in
RT human tissues";
RL J. Recept. Signal Transduct. Res. 18:321-344 (1998).
RN [9]
RP FUNCTION.
RX MEDLINE=98295994; PubMed=9634237;
RX Tachibana K., Hirota S., Iizasa H., Yoshida H., Kawabata K.,
RA Kataoka Y., Kitamura Y., Matsushima K., Yoshida N., Nishikawa S.-I.,
RA Kishimoto T., Nagasawa T.;
RT "The chemokine receptor CXCR4 is essential for vascularization of the
RT gastrointestinal tract";
RL Nature 393:591-594 (1998).
RN [10]
RP FUNCTION.
RX MEDLINE=98295995; PubMed=9634238;
RX Zou Y.-R., Kottmann A.H., Kuroda M., Taniuchi I., Littman D.R.;
RT "Function of the chemokine receptor CXCR4 in hematopoiesis and in
RT cerebellar development";
RL Nature 393:595-599 (1998).
RN [11]
RP DEVELOPMENTAL STAGE.
RX STRAIN=ICR;
RX MEDLINE=99410349; PubMed=10479460;
RA McGrath K.E., Koniski A.D., Maltby K.M., McGann J.K., Palis J.;
RT "Embryonic expression and function of the chemokine SDF-1 and its
RT receptor, CXCR4";
RL Dev. Biol. 213:442-456 (1999).
CC -I- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC signal by increasing the intracellular calcium ions level.
CC Involved in B-cell lymphopoiesis, bone-marrow myelopoiesis and in
CC cardiac ventricular septum formation. Plays also an essential role
CC in vascularization of the gastrointestinal tract, probably by
CC regulating vascular branching and/or remodelling processes in
CC endothelial cells. Involved in cerebellar neuronal layer
CC formation, preventing premature migration of proliferating granule
CC cells from the external granule layer inwards. In the CNS, could
CC mediate hippocampal-neuron survival.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC

CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=CXCR4-B; Synonyms=LESTR-B;
CC IsoId=P70658-1; Sequence=Displayed;
CC Name=CXCR4-A; Synonyms=LESTR-A;
CC IsoId=P70658-2; Sequence=VSP_001891;
CC -I- TISSUE SPECIFICITY: Lymphocytes, macrophages, neutrophils,
CC microglial cells and astrocytes. Found in spleen, thymus, bone
CC marrow, lymph nodes and at lower levels in brain, small
CC intestine, stomach and kidney. CXCR4-A is predominant in all
CC tissues tested.
CC -I- DEVELOPMENTAL STAGE: High expression during embryonic development
CC does not seem to be associated with the differentiation of any
CC particular cell type, but is widely utilized when there is a
CC requirement for cell movement. Frequently associated with less
CC differentiated cell types and down-regulated with subsequent
CC differentiation. Detected in sites with haemopoietic potential:
CC the yolk sac (7.5, 8.5 and 12.5 dpc) and fetal liver (12.5 dpc).
CC During gastrulation, at 7.2 to 7.8 dpc, expressed in the mesoderm
CC and the definitive endoderm. As gastrulation pattern fades (8.5
CC dpc), expression in the mesoderm is down-regulated, while it
CC becomes predominant in neural ectoderm. Endodermal expression is
CC retained in the foregut and later in a subset of foregut
CC derivatives, including the stomach (10.5 dpc), the cystic ducts of
CC the gallbladder and the lung epithelium (12.5 dpc). In neuronal
CC tissue: at 10.5 and 12.5 dpc, expressed in the dorsal root
CC ganglia, in the ventral mantle layer of the spinal cord (or basal
CC plates), in the hindbrain. At 14.5 dpc, expression more tightly
CC confined to the neural epithelium lining the ventricular space and
CC to the external granular layer of the ventral rhombic lip (the
CC developing cerebellum). Expressed in the outpocketing of the
CC diencephalic floor at 10.5 dpc and in the developing thalamus and,
CC to a lesser extent, the developing hypothalamus. At 14.5 dpc,
CC restricted to the region where thalamus and hypothalamus meet.
CC Detected in a discrete band of cells at the edge of the olfactory
CC bulb. In the vascular system: expressed in the endothelium of
CC numerous blood vessels, but not all, at 10.5, 11.5 and 12.5 dpc,
CC such as vitelline/umbilical vessels, cardiac ventricular wall,
CC capillaries, facial vessels and, at 14.5 dpc, in the vasculature
CC of the herniated gut. Expression seems to be associated with
CC expanding vascular networks. In the heart development, expressed
CC at 10.5 dpc in the precursor to the aortopulmonary (AP) septum. At
CC 12.5 dpc, detected in the AP septum at the base of the outflow
CC tract and in the atrioventricular valves. Detected in craniofacial
CC ectoderm from 10.5 to 14.5 dpc. At 10.5 and 11.5 dpc, expressed in
CC the Rathke's pouch.
CC -I- PTM: Sulfated (by similarity).
CC -I- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U59760; AB07725.1; -
CC EMBL; U65580; AAC52953.1; -
CC EMBL; D87747; BAAL3451.1; -
CC EMBL; Z80111; CAB02201.1; -
CC EMBL; Z80112; CAB02202.1; -
CC EMBL; X99581; CAA67893.1; -
CC EMBL; X99582; CAA67894.1; -
CC EMBL; AB000803; BAAL19187.1; -
CC EMBL; BC031685; AAC31665.1; -
CC MG; MG1:109563; CXCR4.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0030334; P:regulation of cell migration; IMP.
CC GO; GO:0042098; P:T-cell proliferation; IMP.
CC InterPro: IPI000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm1.1;
CC PRINTS; PR00237; GPCRHOPOPSN.

Query Match		100.0%;	Score 1867;	DB 1;	Length 359;
Best Local Similarity		100.0%;	Pred. No. 2.2e-115;		
Matches 359;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEPISVITS	DNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIITGIVGNGL	60	
DB	1	MEPISVITS	DNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIITGIVGNGL	60	
QY	61	VILVMGYQKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIITV	120		
DB	61	VILVMGYQKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIITV	120		
QY	121	NYSSVLLIAFISLDYLAIVHATNSQRPKLAERAVVGVWIPALLTIPTDPIFADVS	180		
DB	121	NYSSVLLIAFISLDYLAIVHATNSQRPKLAERAVVGVWIPALLTIPTDPIFADVS	180		
QY	181	QGISQDDRYICDRLYPDSLMWVVFQHIWVGLILPGIVILSCYCIISKLSHSKGHQ	240		
DB	181	QGISQDDRYICDRLYPDSLMWVVFQHIWVGLILPGIVILSCYCIISKLSHSKGHQ	240		
QY	241	KRKALKTTVLLIAFACWLPYVVGISIDSFILLGVIKGCGPESIVHKWISITEALAF	300		
DB	241	KRKALKTTVLLIAFACWLPYVVGISIDSFILLGVIKGCGPESIVHKWISITEALAF	300		
QY	301	HCCNLPILYAFGLAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSSPHSS	359		
DB	301	HCCNLPILYAFGLAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSSPHSS	359		

RESULT 2

CC	RAT	STANDARD;	PRT;	349	AA.
ID	CCR4	RAT			
AC	O0856;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DE	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4)	(Stromal cell-derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived seven transmembrane domain receptor) (LESTR).			
GN	CXCR4	OR CMKAR4.			
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Wistar; TISSUE=Spleen;				
RA	Harrison J.K.; Salafra M.N.;				
RT	"Molecular cloning of rat CXCR4."				
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a signal by increasing the intracellular calcium ions level.				
CC	Involved in B-cell lymphopoiesis, bone marrow myelopoiesis and in cardiac ventricular septum formation. Plays also an essential role in vascularization of the gastrointestinal tract, probably by regulating vascular branching and/or remodelling processes in endothelial cells. Involved in cerebellar neuronal layer formation, preventing premature migration of proliferating granule cells from the external granule layer inwards. In the CNS, could mediate hippocampal-neuron survival (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- TISSUE SPECIFICITY: Expressed in neurons and in astrocytes.				
CC	-1- PTM: Sulfated (By similarity).				
CC	-1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				

CC	EMBL; U90610; AAB50408.1;				
DR	InterPro; IPR000276; GPCR_Rhodpsn.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	PRINTS; PR00237; GPCRHOPOFSN.				
DR	PROSITE; PS0237; G PROTEIN RECP F1 1; 1.				
DR	PROSITE; PS0262; G PROTEIN RECP F1 2; 1.				
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.				
FT	DOMAIN 1 36				
FT	TRANSMEM 37 60				
FT	DOMAIN 61 76				
FT	TRANSMEM 77 96				
FT	DOMAIN 97 107				
FT	TRANSMEM 108 129				
FT	DOMAIN 130 151				
FT	TRANSMEM 152 172				
FT	DOMAIN 173 197				
FT	TRANSMEM 198 217				
FT	DOMAIN 218 237				
FT	TRANSMEM 238 258				
FT	DOMAIN 259 282				
FT	TRANSMEM 283 302				
FT	DOMAIN 303 349				
FT	DISULFID 106 183				
FT	MOD RES 4 4				
FT	MOD RES 18 18				
FT	CARBOHYD 8				
SQ	SEQUENCE 349 AA; 39334 MW; 7E0789A605C60C09 CRC64;				

Query Match

Best Local Similarity	93.4%;	Score 1744.5;	DB 1;	Length 349;
Matches 338;	Conservative 4;	Mismatches 7;	Indels 5;	Gaps 1;

QY	6	VSIYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIITGIVGNGLVLM	65
DB	1	MEIYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTIYFIITGIVGNGLVLM	60
QY	66	GQYKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIITVNL	125
DB	61	GQYKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKFLCKAVHIITVNL	120
QY	126	VLLIAFISLDYLAIVHATNSQRPKLAERAVVGVWIPALLTIPTDPIFADVSQGD	185
DB	121	VLLIAFISLDYLAIVHATNSQRPKLAERAVVGVWIPALLTIPTDPIFADVSQGD	175
QY	186	QGDRIYICDRLYPDSLMWVVFQHIWVGLILPGIVILSCYCIISKLSHSKGHQK	245
DB	176	QGDRIYICDRLYPDSLMWVVFQHIWVGLILPGIVILSCYCIISKLSHSKGHQK	235
QY	246	KTTVLLIAFACWLPYVVGISIDSFILLGVIKGCGPESIVHKWISITEALAFH	305
DB	236	KTTVLLIAFACWLPYVVGISIDSFILLGVIKGCGPESIVHKWISITEALAFH	295
QY	306	PILYAFGLAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSSPHSS	359
DB	296	PILYAFGLAKFKSSAQHALNSMRGSSLKILSKGRGHSVSTESSSPHSS	349

RESULT 3

CC	PAPAN	STANDARD;	PRT;	352	AA.
ID	CCR4	PAPAN			
AC	P56431;				
DT	15-JUL-1998	(Rel. 36, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4)	(SDF-1 receptor) (stromal cell-derived factor 1 receptor) (Fusin).			
GN	CXCR4.				
OS	Papio anubis (Olive baboon).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;				
OC	Cercopitheciinae; Papio.				
OX	NCBI_TaxID=9555;				

QY 126 VLIATFISLDYLAIVHATNSQRPKLAKEAVYGVWIPALLITIPDIPADVSQGDIS 185
 Db 124 VLIATFISLDYLAIVHATNSQRPKLAKEAVYGVWIPALLITIPDIPFASV-----S 178
 QY 186 QGDRYICDRYPDSLMVVFQFHIMVGLPGIVILSCYCIITIKLSHSGHQKRAL 245
 Db 179 EADDRYICDRYPDNLWVVFQFHIMVGLPGIVILSCYCIITIKLSHSGHQKRAL 238
 QY 246 KTVVILAFACWLPYVYGISDSTFLGVKOGDFESIVHKWISITEALAFHCCLN 305
 Db 239 KTVVILAFACWLPYVYGISDSTFLGVKOGDFESIVHKWISITEALGFHCCLN 298
 QY 306 PLYAFILGAKFKSAQHALMSRSGSLKILSKGKGHSHSVSTESSESPHSS 359
 Db 299 PLYAFILGAKFKSAQHALTSVSRGSLKILSKGKGHSHSVSTESSESPHSS 352

RESULT 5
 CCR4 HUMAN STANDARD; PRT; 352 AA.
 AC P30951; O60835; P56438; Q9UKN2;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (Stromal cell-
 DE derived factor 1 receptor) (SDF-1 receptor) (Fusin) (Leukocyte-derived
 DE seven transmembrane domain receptor) (LESTR) (LCR1) (PB22) (NPYRL)
 DE (HM89) (CD184 antigen).
 GN CXCR4.
 OS Homo sapiens (Human), and
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606, 9598;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Lung;
 RX MEDLINE=93319623; PubMed=8329116;
 RA Herzog H., Hort Y.J., Shine J., Selbie L.A.;
 RT "Molecular cloning, characterization, and localization of the human
 RT homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
 RT and activation.";
 RL DNA Cell Biol. 12:465-471(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Fetal brain;
 RX MEDLINE=94052833; PubMed=8234909;
 RA Jazin E.E., Yoo H., Blomqvist A.G., Yee F., Weng G., Walker M.W.,
 RA Salton J., Larhammar D., Wahlestedt C.R.;
 RT "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
 RT human homologue, confers neither NPY binding sites nor NPY
 RT responsiveness on transfected cells.";
 RL Regul. Pept. 47:247-258(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Fetal spleen;
 RX MEDLINE=93315164; PubMed=8325644;
 RA Federapfel B., Melhado I.G., Durcan A.M., Delaney A.D.,
 RA Schappert K.T., Clark-Lewis I., Jirik F.R.;
 RT "Molecular cloning of the cDNA and chromosomal localization of the
 RT gene for a putative seven-transmembrane segment (7-TMS) receptor
 RT isolated from human spleen.";
 RL Genomics 16:707-712(1993).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Monocytes;
 RX MEDLINE=94103215; PubMed=8276799;
 RA Loetscher M., Geiser T., O'Reilly T., Zwaalen R., Baggiolini M.,
 RA Moser B.;
 RT "Cloning of a human seven-transmembrane domain receptor, LESTR, that
 RT is highly expressed in leukocytes.";
 RL J. Biol. Chem. 269:232-237(1994).

RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Peripheral blood leukocytes;
 RX MEDLINE=94092629; PubMed=7505609;
 RA Nomura H., Nielsen B.W., Matsushima K.;
 RT "Molecular cloning of cDNAs encoding a LD78 receptor and putative
 RT leukocyte chemotactic peptide receptors.";
 RL Int. Immunol. 5:1239-1249(1993).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION OF ITS HIV-1
 RC CORECEPTOR FUNCTION.
 RX SPECIES=Human;
 RX MEDLINE=96211947; PubMed=8629022;
 RA Feng Y., Broder C.C., Kennedy P.E., Berger E.A.;
 RT "HIV-1 entry cofactor: functional cDNA cloning of a seven-
 RT transmembrane, G protein-coupled receptor.";
 RL Science 272:872-877(1996).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Peripheral blood leukocytes;
 RX MEDLINE=98136183; PubMed=9468539;
 RA Wegner S.A., Ehrenberg P.K., Chang G., Dayhoff D.E., Sleeker A.L.,
 RA Michael N.L.;
 RT "Genomic organization and functional characterization of the chemokine
 RT receptor CXCR4, a major entry co-receptor for human immunodeficiency
 RT virus type 1.";
 RL J. Biol. Chem. 273:4754-4760(1998).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human;
 RX MEDLINE=98258970; PubMed=9599023;
 RA Caruz A., Sansom M., Alonso J.M., Alcamí J., Baleux F.,
 RA Virielizier J.L., Parmentier M., Arenzana-Seisdedos F.;
 RT "Genomic organization and promoter characterization of human CXCR4
 RT gene.";
 RL FEBS Lett. 426:271-278(1998).
 RN [9]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human;
 RX MEDLINE=99408510; PubMed=10480633;
 RA Xiao L., Weiss S.H., Qari S.H., Rudolph D., Zhao C., Denny T.N.,
 RA Hodge T., Lal R.B.;
 RT "Partial resistance to infection by RSX4 primary HIV type 1 isolates
 RT in an exposed-uninfected individual homozygous for CCR5 32-base pair
 RT deletion.";
 RL AIDS Res. Hum. Retroviruses 15:1201-1208(1999).
 RN [10]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Peripheral blood lymphocytes;
 RX MEDLINE=99095114; PubMed=9879064;
 RA Frodl R., Gierschik P., Moepps B.;
 RT "Genomic organization and expression of the CXCR4 gene in mouse and
 RT man: absence of a splice variant corresponding to mouse CXCR4-B in
 RT human tissues.";
 RL J. Recept. Signal Transduct. Res. 18:321-344(1998).
 RN [11]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC SPECIES=Human; TISSUE=Neutrophils;
 RX MEDLINE=99384048; PubMed=10452968;
 RA Gupta S.K., Pillarsetti K.;
 RT "CXCR4-10: molecular cloning and functional expression of a novel
 RT human CXCR4 splice variant.";
 RL J. Immunol. 163:2368-2372(1999).
 RN [12]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Lung;
 RA Warren C.N., Aronstam R.S., Sharma S.V.;
 RT "cDNA clones of human proteins involved in signal transduction
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC SPECIES=Human; TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Farney J., Helton E., Kettner M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.C., Truchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [14]
RP SEQUENCE FROM N.A.
RC SPECIES=P.troglodytes;
RX MEDLINE=98090115; PubMed=94302050;
RA Pretet J.-L., Zerbib A.C., Girard M., Guillet J.-G., Butor C.;
RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";
RL AIDS Res. Hum. Retroviruses 13:1583-1587 (1997).
RN [15]
RP SULFATION.
RC SPECIES=Human;
RX MEDLINE=99189752; PubMed=10089882;
RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
RA Gerard N.P., Gerard C., Sodroski J., Choe H.;
RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
RT entry.";
RL Cell 96:667-676 (1999).
RN [16]
RP FUNCTION.
RX MEDLINE=96351077; PubMed=8752280;
RA Bleul C.C., Farzan M., Choe H., Farolan C., Clark-Lewis I.,
RA Sodroski J., Springer T.A.;
RT "The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and
RT blocks HIV-1 entry.";
RL Nature 382:829-833 (1996).
RN [17]
RP FUNCTION.
RX MEDLINE=96351078; PubMed=8752281;
RA Oberlin E., Amara A., Bachevalier J., Bessia C., Virelizier J.-L.,
RA Arenzana-Seisdedos F., Schwartz O., Heard J.-M., Clark-Lewis I.,
RA Legler D.F., Loetscher M., Baggiolini M., Moser B.;
RT "The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents
RT infection by T-cell-line-adapted HIV-1.";
RL Nature 382:833-835 (1996).
RN [18]
RP ERRATUM.
RX Oberlin E., Amara A., Bachevalier J., Bessia C., Virelizier J.-L.,
RA Arenzana-Seisdedos F., Schwartz O., Heard J.-M., Clark-Lewis I.,
RA Legler D.F., Loetscher M., Baggiolini M., Moser B.;
RL Nature 384:288-288 (1996).
RN [19]
RP CHARACTERIZATION OF ITS HIV-1 CORECEPTOR FUNCTION.
RX MEDLINE=97002453; PubMed=8849450;
RA Lapham C.K., Ouyang J., Chandrasekhar B., Nguyen N.Y., Dimitrov D.S.,
RA Golding H.;
RT "Evidence for cell-surface association between fusin and the CD4-gp120
RT complex in human cell lines.";
RL Science 274:602-605 (1996).
RN [20]
RP CHARACTERIZATION OF ITS HIV-2 RECEPTOR FUNCTION.
RX MEDLINE=97083584; PubMed=89259542;
RA Endres M.J., Clapham P.R., Marsh M., Ahuja M., Turner J.D.,
RA McKnight A., Thomas J.F., Stoeckenau-Haggarty B., Choe S., Vance P.J.,
RA Wells T.N.C., Power C.A., Sutterwala S.S., Doms R.W., Landau N.R.,

RA Hoxie J.A.;
Query Match 89.9%; Score 1678.5; DB 1; Length 352;
Best Local Similarity 90.1%; Pred No. 4.5e-103;
Matches 319; Conservative 19; Mismatches 11; Indels 5; Gaps 1;
QY 6 VSIYTSNYSSEVGSDYDSNKEPCFRBNVHFNRIPLTYFIIFLTGIVNGLVILVM 65
DB 4 ISIVTSDNYTEEMSGSDYDSNKEPCFRBNVHFNRIPLTYFIIFLTGIVNGLVILVM 63
QY 66 GYQKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIYTVNLYSS 125
DB 64 GYQKLRMTDKYRLHLSVADLLFVITLPFWAVDAMADWYFGKPLCKAVHIYTVNLYSS 123
QY 126 VLILAFISLDRLAIVHATNSQRPKLLAEKAVYGVWIPALLTIPDIFADVQGDIS 185
DB 124 VLILAFISLDRLAIVHATNSQRPKLLAEKAVYGVWIPALLTIPDIFANV-----S 178
QY 186 QGDDYICDRLYPDLSLWVWVFOHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 245
DB 179 EADRYICDRFPNDLWVWVFOHIMVGLILPGIVILSCYCIISKLSHSGHOKRKAL 238
QY 246 KTVTILILAFFACWLPYVVGISIDSIFILLGVTKQCFESIVHKWISITEALAFPHCLN 305
DB 239 KTVTILILAFFACWLPYVVGISIDSIFILLGVTKQCFESIVHKWISITEALAFPHCLN 298
QY 306 PIIYAFLGKFKXSQAHALNSMRGSSUKILSKGRGGHSSVSTESSESSPHSS 359
DB 299 PIIYAFLGKFKXSQAHALNSMRGSSUKILSKGRGGHSSVSTESSESSPHSS 352
RESULT 6
CCR4_MACMU STANDARD; PRT; 352 AA.
ID CCR4_MACMU STANDARD; PRT; 352 AA.
AC P79394; O02745; O46428;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (SDF-1 receptor)
DE (Stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN CXCR4.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Indian macaque;
RX MEDLINE=97213934; PubMed=9060623;
RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
RT "Genetically divergent strains of simian immunodeficiency virus use
RT CCR5 as a coreceptor for entry.";
RL J. Virol. 71:2705-2714 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharron M., Samson M., Lu Z.-H., Clements J.E., Murphy-Corb M.,
RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT "Differential utilization of CCR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Chinese;
RX MEDLINE=98252393; PubMed=9591719;
RA Pretet J.-L., Guillet J.-G., Butor C.;
RT "New widespread CXCR4 allele in rhesus macaques does not predict
RT subtypes or clinical evolution.";
RL AIDS Res. Hum. Retroviruses 14:639-641 (1998).
CC -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC signal by increasing the intracellular calcium ions level.


```

RESULT 7
CCR4 CCR4 STANDARD; PRT; 352 AA.
AC 062747;
AC 15-DEC-1998 (Rel. 37, Created)
AC 15-DEC-1998 (Rel. 37, Last sequence update)
AC 16-OCT-2001 (Rel. 40, Last annotation update)
DT C-X-C chemokine receptor type 4 (CXCR-4) (SDF-1 receptor)
DE (Stromal cell-derived factor 1 receptor) (Fusin) (UESIR).
DE CXCR4.
GN Cercocebus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercocebus.
OC NCBI_TaxID=9531;
OX [1]_RN
RX SEQUENCE FROM N.A.
RP MEDLINE=98321155; PubMed=9656999;
RA Chen Z., Gettie A., Ho D.D., Marx P.A.;
RA "Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys
RT naturally infected in west Africa: a comparison of coreceptor usage
RT of primary SIVsm, HIV-2, and SIVmac.";
RT Virology 246:113-124(1998).
CC -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC signal by increasing the intracellular calcium ions level.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF051906; AAC39834.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G PROTEIN RECP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
KW DOMAIN 1 39 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 40 63 1 (POTENTIAL).
FT DOMAIN 64 79 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 80 99 2 (POTENTIAL).
FT DOMAIN 100 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 132 3 (POTENTIAL).
FT DOMAIN 133 154 4 (POTENTIAL).
FT TRANSMEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 200 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 201 220 6 (POTENTIAL).
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 7 (POTENTIAL).
FT DOMAIN 262 285 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 286 305 8 (POTENTIAL).
FT DOMAIN 306 352 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 21 21 SULFATION (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFD 109 186 BY SIMILARITY.
FT SEQUENCE 352 AA; 39648 MW; 4027432B8032F87E CRC64;
Query Match 89.4%; Score 1668.5; DB 1; Length 352;
Best Local Similarity 89.5%; Pred No. 28-102;
Matches 317; Conservative 19; Mismatches 13; Indels 5; Gaps 1;
QY 6 VSIYTDNYSYSEVSGSDYDSNKEPCFRDENVFNFRIFLTIIYFIIFLTGIVGNGLVILWM 65
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 IGIYTDNYSYSEVSGSDYDSNKEPCFRDENVFNFRIFLTIIYFIIFLTGIVGNGLVILWM 63

```

Db 64 GYQKLRSMYDYLRLSLVADLLFVITLFPWAVDANWYFGNPLCKAVHVIYTVNLYSS 123
QY 126 VLIILAFSLDRYLALVHATNSORERKLLAEKAVVGVWIPALLLTIPDFADVSCQDIS 185
Db 124 VLIILAFSLDRYLALVHATNSORERKLLAEKAVVGVWIPALLLTIPDFADVSCQDIS 178
QY 186 QGDRYICDRILYPSLWVWVVFQOHIMVGLIPGIVILSCYCIISKLKSHKQKRRKAL 245
Db 179 EADDRFICDRFPNDLWVWVVFQOHIMVGLIPGIVILSCYCIISKLKSHKQKRRKAL 238
QY 246 KTVILLILAFACWLPYVIGISIDSFILLGVIKOGCDFESIVHKWISITALEAFFHCCLN 305
Db 239 KTVILLILAFACWLPYVIGISIDSFILLGVIKOGCDFESIVHKWISITALEAFFHCCLN 298
QY 306 PILYAFILGAKFKSAQHALNSMRGSSILKLSKGRGSHSVSTESSESSSFHSS 359
Db 299 PILYAFILGAKFKSAQHALNSMRGSSILKLSKGRGSHSVSTESSESSSFHSS 352

RESULT 8
CCR4_FELCA STANDARD; PRT; 353 AA.
ID CCR4_FELCA STANDARD; PRT; 353 AA.
AC P56458; 002700; P79172;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (SDF-1 receptor)
DE (stromal cell-derived factor 1 receptor) (Fusin) (LESTR).
GN CXCR4.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97404646; PubMed=9261358;
RA Willett B.J., Picard L., Hosie M.J., Turner J.D., Adema K.,
RA Clapham P.R.;
RT "Shared usage of the chemokine receptor CXCR4 by the feline and human
immunodeficiency viruses";
RL J. Virol. 71:6407-6415(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Willett B.J.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lerner D.L., Elder J.H.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=99399239; PubMed=10470253;
RA Kovacs E.M., Baxter G.D., Robinson W.F.;
RT "Feline peripheral blood mononuclear cells express message for both
CXCR4 and CC type chemokines";
RL Arch. Virol. 144:273-285(1999).
CC -!- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
signal by increasing the intracellular calcium ions level.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U63558; AAC48952.1; -.
CC EMBL; U92795; AAB51765.1; -.

DR EMBL; AJ009816; CAA08839.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECP F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 40
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 64
FT DOMAIN 65 80
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 100
FT DOMAIN 101 111
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 112 133
FT DOMAIN 134 155
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176
FT DOMAIN 177 201
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 202 221
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 242 262
FT DOMAIN 263 286
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 287 306
FT DOMAIN 307 353
FT CYTOPLASMIC (POTENTIAL).
FT MOD RES 22 22
FT CARBOHYD 11 11
FT DISULFID 110 187
FT CONFLICT 67 67
FT CONFLICT 263 263
SQ SEQUENCE 353 AA; 39935 MW; EA2BD46068A6C05B CRC64;
Query Match 88.0%; Score 1643; DB 1; Length 353;
Best Local Similarity 89.0%; Pred. No. 9.4e-101;
Matches 314; Conservative 18; Mismatches 15; Indels 6; Gaps 2;
QY 8 IYTDNYSSE-EVSGDYSNKPFCFRDENVHNRIPLTYFIIFLTGIVGNGLVILVWG 66
Db 6 IYTDNYSSE-EVSGDYSNKPFCFRDENVHNRIPLTYFIIFLTGIVGNGLVILVWG 65
QY 67 YQKLRSMYDYLRLSLVADLLFVITLFPWAVDANWYFGNPLCKAVHVIYTVNLYSSV 126
Db 66 YQKLRSMYDYLRLSLVADLLFVITLFPWAVDANWYFGNPLCKAVHVIYTVNLYSSV 125
QY 127 LILAFISLDRYLALVHATNSORERKLLAEKAVVGVWIPALLLTIPDFADVSCQDISQ 186
Db 126 LILAFISLDRYLALVHATNSORERKLLAEKAVVGVWIPALLLTIPDFADVSCQDISQ 180
QY 187 GDDRYICDRILYPSLWVWVVFQOHIMVGLIPGIVILSCYCIISKLKSHKQKRRKAL 246
Db 181 AGRYICDRILYPSLWVWVVFQOHIMVGLIPGIVILSCYCIISKLKSHKQKRRKAL 240
QY 247 TTVILLILAFACWLPYVIGISIDSFILLGVIKOGCDFESIVHKWISITALEAFFHCCLNP 306
Db 241 TTVILLILAFACWLPYVIGISIDSFILLGVIKOGCDFESIVHKWISITALEAFFHCCLNP 300
QY 307 ILVAFILGAKFKSAQHALNSMRGSSILKLSKGRGSHSVSTESSESSSFHSS 359
Db 301 ILVAFILGAKFKSAQHALNSMRGSSILKLSKGRGSHSVSTESSESSSFHSS 353

RESULT 9
CCR4_BOVIN STANDARD; PRT; 353 AA.
ID CCR4_BOVIN STANDARD; PRT; 353 AA.
AC P5930;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (SDF-1 receptor)
DE (stromal cell-derived factor 1 receptor) (Fusin) (Leukocyte-derived
seven transmembrane domain receptor) (LESTR) (LCR1).
GN CXCR4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Locus coeruleus;
RX MEDLINE=92100053; PubMed=1661837;
RA Rimaland J., Xin W., Sweetnam P., Saijoh K., Nestler E.J., Duman R.S.;
RT "Sequence and expression of a neuropeptide Y receptor cDNA.";
RL Mol. Pharmacol. 40:869-875 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Werling D.;
RT "Role of chemokines in respiratory syncytial virus infection.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SHOWS THAT IT IS NOT A NPV3-R.
RX MEDLINE=94052833; PubMed=8234909;
RA Jazin E., Yoo H., Blomqvist G., Yee F., Weng G., Walker M.W.,
RA Salton J., Larhammar D., Wahlstedt C.R.;
RT "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
RT human homologue, confers neither NPY binding sites nor NPY
RT responsiveness on transfected cells.";
RL Regul. Pept. 47:247-258 (1993).
CC -1- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC signal by increasing the intracellular calcium ions level.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Brain, heart, kidney, lung and liver.
CC -1- PTM: Sulfated (By similarity).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- CAUTION: Was originally (Ref.1) thought to be a receptor for
CC neuropeptide Y type 3 (NPY3-R).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M86739; -; NOT ANNOTATED CDS.
DR EMBL; AF399642; AAK94452.1; -;
DR PIR; S28787; S28787;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOSN.
DR PROSITE; PS00237; G PROTEIN RECEPT FL 1; 1.
DR PROSITE; PS0262; G PROTEIN RECEPT FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 64
FT DOMAIN 65 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 100
FT DOMAIN 101 111 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 112 133
FT DOMAIN 134 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176
FT DOMAIN 177 201 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 202 221
FT DOMAIN 222 241 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 242 262
FT DOMAIN 263 286 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 287 306
FT DOMAIN 307 353 CYTOPLASMIC (POTENTIAL).
FT MOD RES 32 22 SULFATION (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLNAC. . .) (POTENTIAL).
FT DISULFID 110 187 BY SIMILARITY.
SQ SEQUENCE 353 AA; 39938 MW; 42FFE5BC754505E CRC64;
Query Match 87.4%; Score 1631; DB 1; Length 353;
Best Local Similarity 86.8%; Pred. No. 5,7e-100;
Matches 308; Conservative 26; Mismatches 15; Indels 6; Gaps 2;
QY 6 VSIYTSNYSSE-EVSGDYSNKEPCRDENHFNRIPLFTYIFLITGIVGNGLVILV 64

Db 4 TRIFTSNEDDLSGSDYDSMKPCPREENAHENRIFLPTVYSIFLTGIVGNGLVILV 63
QY 65 NGYQKLRSMQDKYRLHLVADLLFVITLTPWAVDAMADWTFGKFLCAVHIITVNLVS 124
Db 64 NGYQKLRSMQDKYRLHLVADLLFVITLTPWAVDAMADWTFGKFLCAVHIITVNLVS 123
QY 125 SVLLIAPISLDRLYLAIHVATNSQRPKLLAEKAVVGVWIPALLLTTPDFADVSQGD 184
Db 124 SVLLIAPISLDRLYLAIHVATNSQRPKLLAEKAVVGVWIPALLLTTPDLFA----DI 178
QY 185 SQGDDRYICDRLYPDSLMVVFQFHIMVGLILPGIIVILSCYCIISKLSHSKGQKKA 244
Db 179 KEVDERYICDRFYPDSLWLWVVFQFHIVGGLLPGIIVILSCYCIISKLSHSKGQKKA 238
QY 245 LKTTVILLIAPFACWLPVYVIGISIDSFLLGVIKOGCDFESIVHKWISITALLAFHCL 304
Db 239 LKTTVILLIAPFACWLPVYVIGISIDSFLLGVIKOGCDFESIVHKWISITALLAFHCL 298
QY 305 NPILYAFILGAKFKSSAQHALNSMRGSSKLKLSKGGKGGHSSVSTESSESSFHSS 359
Db 299 NPILYAFILGAKFKTSQAHLTVSRRGSSKLKLSKGGKGGHSSVSTESSESSFHSS 353

CC RESULT 10
CC CCR4 SHEEP
CC ID CCR4 SHEEP STANDARD; PRT; 192 AA.
AC Q28553;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE C-X-C chemokine receptor type 4 (CXCR-4) (CXCR-4) (SDF-1 receptor)
DE (Stromal cell-derived factor 1 receptor) (Fusin) (Leukocyte-derived
DE seven transmembrane domain receptor) (LESTR) (Fragment).
GN CXCR4 OR LESTR.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OC NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Dyer C.J., Matteri R.L., Keisler D.H.;
RT "Development of an ovine Y3 cDNA and expression of the Y3 receptor
RT mRNA in the ovine hypothalamus and pituitary.";
RL Abstr. - Soc. Neurosci. 21:1890-1890 (1995).
CC -1- FUNCTION: Receptor for the C-X-C chemokine SDF-1. Transduces a
CC signal by increasing the intracellular calcium ions level.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: Sulfated (By similarity).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- CAUTION: Was originally (Ref.1) thought to be a receptor for
CC neuropeptide Y type 3 (NPY3-R).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U38942; AAA81347.1; -;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOSN.
DR PROSITE; PS00237; G PROTEIN RECEPT FL 1; 1.
DR PROSITE; PS0262; G PROTEIN RECEPT FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 29 EXTRACELLULAR (POTENTIAL).
FT NON TER <1 29
FT DOMAIN 30 53
FT TRANSMEM 30 53

FT DOMAIN 54 69 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 70 89 2 (POTENTIAL).
FT DOMAIN 90 100 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 101 122 3 (POTENTIAL).
FT DOMAIN 123 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 165 4 (POTENTIAL).
FT DOMAIN 166 190 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 191 >192 5 (POTENTIAL).
FT DISULFID 99 176 BY SIMILARITY.
FT MOD RES 11 11 SULFATION (POTENTIAL).
FT NON_TER 192 192
SQ SEQUENCE 192 AA; 22178 MW; ABBCEP30352BD98 CRC64;

Query Match 47.4%; Score 885.5; DB 1; Length 192;
Best Local Similarity 83.1%; Pred. No. 1.7e-51;
Matches 167; Conservative 14; Mismatches 11; Indels 9; Gaps 2;

QY 9 YTSNDYSEVSGDYDSNKEPCFRDENVHFRNIFLPTIYFIITGIVGNGLVILVMGYQ 68
DB 1 YTED----DLGSDYDSNKEPCFRENAHFRNIFLPTIYFIITGIVGNGLVILVMGYQ 56
QY 69 KKLRSMTKYRLHLSVADLLFVITLPPFWADAMADWYFGKFLCKAVHIIYTNLYSSVLI 128
DB 57 KKLRSMTKYRLHLSVADLLFVITLPPFWADAMADWYFGKFLCKAVHIIYTNLYSSVLI 116
QY 129 LAFISLDVLAIVHATNSORPEKLAEKAVYGVWIPALLTIPDFADVSQGDISQGD 188
DB 117 LAFISLDVLAIVHATNSORPEKLAEKAVYGVWIPALLTIPDFADVSQGDISQGD 171

QY 189 DRYICDRYPDSLWVVFQFQ 209
DB 172 ERYICDRFYPDSLWVVFQFQ 192

RESULT 11
CCR3 HUMAN
ID CCR3 HUMAN STANDARD; PRT; 368 AA.
AC P49682; O15185; Q9P274; Q9P275;
DT 01-REB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE C-X-C chemokine receptor type 3 (CCR3) (CCR-3) (CCR-L2) (CD183 antigen).
DE CCR3 OR GPR9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Blood;
RX MEDLINE=97188912; PubMed=9064356;
RA Clark-Lewis I., Baggiolini M., Moser B.,
RT "Chemokine receptor specific for IP10 and MIG: structure, function,
RL and expression in activated T-lymphocytes.";
RL J. Exp. Med. 184:963-969(1996).
[2]
SEQUENCE FROM N.A.
RA Gutierrez J., Varona R., Zaballos A., Lind P., Marquez G.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Warren C.N., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
RL sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RP TISSUE=Brain, Lung, and Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallick D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
SEQUENCE OF 5-368 FROM N.A.
RX MEDLINE=96115583; PubMed=8666380;
RA Marchese A., Heiber M., Nguyen T., Heng H.H.Q., Saldivia V.R.,
RA Cheng R., Murphy P.M., Tsui L.-C., Shi X., Gregor P., George S.R.,
RA O'Dowd B.F., Docherty J.M.;
RT "Cloning and chromosomal mapping of three novel genes, GPR9, GPR10,
RT Y, and somatostatin receptors.";
RL Genomics 29:335-344(1995).
[6]
SEQUENCE OF 278-368 FROM N.A., AND VARIANTS GLN-292 AND THR-363.
RX MEDLINE=21040285; PubMed=11196695;
RA Kato H., Tsuchiya N., Tokunaga K.;
RT "Single nucleotide polymorphisms in the coding regions of human CX-
RT chemokine receptors CXCR1, CXCR2 and CXCR3.";
RL Genes Immun. 1:330-337(2000).
[7]
LIGAND BINDING.
RP TISSUE=Fetal astrocytes;
RX MEDLINE=98290735; PubMed=9625760;
RA Cole K.E., Strick C.A., Paradis T.J., Ogborne K.T., Loetscher M.,
RA Gladue R.P., Lin W., Boyd J.G., Moser B., Wood D.E., Sahagan B.G.,
RA Neote K.;
RT "Interferon-inducible T cell alpha chemoattractant (I-TAC): A novel
RT non-EUR CX chemokine with potent activity on activated T cells
RT through selective high affinity binding to CXCR3.";
J. Exp. Med. 187:2009-2021(1998).
CC -1- FUNCTION: Receptor for SCYB9/MIG, SCYB10/INP10 and SCYB11/ITAC.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -1- DATABASE: NAME=PROW; NOTE=PROW 1:78-83(2000).
WWW="http://www.ncbi.nlm.nih.gov/prov/guide/650534941_g.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X95876; CAA65126.1; -
EMBL; Z9783; CAB02143.1; ALT INIT.
EMBL; AY242128; AA092295.1; -
EMBL; BC034403; AA034403.1; -
EMBL; U32674; AAC50505.1; -
EMBL; AB032735; BAA92297.1; -
EMBL; AB032736; BAA92298.1; -
Genew; HGNC:4540; CXCR3.
MIM; 600894; -
DR GO; GO:0005737; Cytoplasm; TAS
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0004950; P: chemokine receptor activity; TAS.
DR GO; GO:0006960; P: antimicrobial humoral response (sensu Inver. .; TAS.

DR GO: GO:0007155; P:cell adhesion; TAS.
DR GO: GO:0006928; P:cell motility; TAS.
DR GO: GO:0006935; P:chemotaxis; TAS.
DR GO: GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR InterPro: IPR004070; CXC 3 receptor.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR01532; CXCHMKINER3.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE: PS00262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Antigen;
KW Polymorphism.
FT DOMAIN 1 53 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 54 80 1 (POTENTIAL).
FT DOMAIN 81 89 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 90 110 2 (POTENTIAL).
FT DOMAIN 111 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 147 3 (POTENTIAL).
FT DOMAIN 148 169 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 170 189 4 (POTENTIAL).
FT DOMAIN 190 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 233 5 (POTENTIAL).
FT DOMAIN 234 255 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 256 277 6 (POTENTIAL).
FT DOMAIN 278 298 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 299 321 7 (POTENTIAL).
FT DOMAIN 322 368 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 203 BY SIMILARITY.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 292 292 R -> Q.
FT VARIANT 363 363 A -> T.
FT CONFLICT 75 75 /FTID=VAR_016241.
FT SEQUENCE 368 AA; 40659 MW; F08A3B4B2BBAD04 CRC64;

Query Match 31.3%; Score 584; DB 1; Length 368;
Best Local Similarity 35.2%; Pred. No. 1.5e-31;
Matches 124; Conservative 67; Mismatches 139; Indels 22; Gaps 5;

QY 12 DNYSEVSGSDYDN-----KEPCFRDENVHNRIFLPTIYFIPLTGIVNGVLVIMGYQKLRMT 67
DB 21 ENFSSSYDGENESDSCCTSPCPQDFSLNFDRAFLPALYSLLFLGLGNGAVALLS 80
QY 68 OKLRSMTDKYRLHLSVADLLVITLPPFVAVDAMADWYFGKFLCKAVHIIYTNLYSSVL 127
DB 81 RTALSTDTFLHLAVADLLVITLPLWAVDAVQVWFGSLGCKVAGALFNINFFAGAL 140
QY 128 ILAFISLDYLAIVHATNSQRPKLLAEKAVYGVVWIPALLTIPDFADVSQGDISI 187
DB 141 LLACISPDYLVNIHATQYLRGPPARVTLTCLAVWGLCLLFPALPDFILS-----AHH 194
QY 188 DDRY---ICDRLYDLSLWVVFQFHLMVGLLIPGVILSCYCIILSKLSHGKHKRKA 244
DB 195 DERLNATHCQNFQ-QGRVLRVQLVAVGLFLLWVAYCAHILAVLLVSRGRLRA 253
QY 245 LKTVILILAFFACWLPYVYGISDIFLLGVILKQGDFFESIVHKWISITELAFPHCL 304
DB 254 MRLVVVVVAFALCTPYHLVVLVDILMDLGALRNGRESRDVAKSVTSGLGYMHCC 313
QY 305 NPILYAFGLAKFKSQAHALNSMRGSLKILSKRGKGGHSSVTSSESSP 356
DB 314 NPLLYAFYGVKFRERMMWLL-----LRLGCPNQRGLQRPSSRRDSSW 357

RESULT 12
CCR3_MOUSE
ID CCR3_MOUSE STANDARD; PRT; 367 AA.
AC O88410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
C-X-C chemokine receptor type 3 (CXC-R3) (CXCR-3).
CXCR3 OR CMKAR3.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=98318636; PubMed=9653165;
Soto H., Wang W., Strieter R.M., Copeland N.G., Gilbert D.J.,
Jenkins N.A., Hedrick J., Zlotnik A.;
"The CC chemokine 6CKine binds the CXC chemokine receptor CXCR3.";
Proc. Natl. Acad. Sci. U.S.A. 95:8205-8210(1998).
-!- FUNCTION: Receptor for SCYB9/MIG, SCYB10/INP10 and SCYB11/ITAC (By
similarity). Binds to SCYB21.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; AF045146; AAC40163.1; .
MG; MG1.1277207; CXC-R3.
InterPro; IPR004070; CXC 3 receptor.
InterPro; IPR000276; GPCR_Rhodopsn.
Pfam; PF00001; 7tm_1; 1.
PRINTS; PR01532; CXCHMKINER3.
PRINTS; PR00237; GPCR_Rhodopsn.
PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 79 1 (POTENTIAL).
FT DOMAIN 80 88 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 89 109 2 (POTENTIAL).
FT DOMAIN 110 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 146 3 (POTENTIAL).
FT DOMAIN 147 168 4 (POTENTIAL).
FT TRANSMEM 169 188 5 (POTENTIAL).
FT DOMAIN 189 211 6 (POTENTIAL).
FT TRANSMEM 212 232 7 (POTENTIAL).
FT DOMAIN 233 254 8 (POTENTIAL).
FT TRANSMEM 255 276 9 (POTENTIAL).
FT DOMAIN 277 297 10 (POTENTIAL).
FT TRANSMEM 298 320 11 (POTENTIAL).
FT DOMAIN 321 367 BY SIMILARITY.
FT DISULFID 123 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 367 AA; 41016 MW; 029FBE778E3CD4EA CRC64;

Query Match 31.1%; Score 581.5; DB 1; Length 367;
Best Local Similarity 35.2%; Pred. No. 2.2e-31;
Matches 122; Conservative 70; Mismatches 130; Indels 25; Gaps 7;

QY 17 EVSGSDYD-SNKEPCFRDENVHNRIFLPTIYFIPLTGIVNGVLVIMGYQKLRMT 75
DB 28 DYGENESDSDSPCPQDFSLNFDRAFLPALYSLLFLGLGNGAVALLSQTALST 87
QY 76 DKYRLHLSVADLLVITLPPFVAVDAMADWYFGKFLCKAVHIIYTNLYSSVLILAFSLD 135
DB 88 DTFLHLAVADLLVITLPLWAVDAVQVWFGSLGCKVAGALFNINFFAGALFACISFD 147
QY 136 RYLAIVHATNSQRPKLLAEKAVYGVVWIPALLTIPDFADVSQGDISIQQDDRYI- 192
DB 148 RYLSIVHATQIYRDRPRVRLTCLV--VWGLCLLFPALPDFIY-----LSANYDQRLN 198

QY 193 ---CDRLPDSLWVVVFQFHIMVGLILPGIVILSCYIIILKSLSHSGHGRKALKTTV 249
Db 199 ATHQVNFQ-QVGRFALRVQLVAGFLPLLVAYCYAHILAVILVSRGQFRFAMRLW 257
QY 250 ILILFAFFACWLEYYVIGISIDSFILGLVKGQGFESIVHKWISITEALAFPHCCINPLY 309
Db 258 VVVAFAVCTYHLVWLVDILMDVGLARNGRSHVDVAKSVYSGVYHCCINPLY 317
QY 310 AFLGAKFKSSAQHALNSMRGSLKILSGKGGHSSVSTSESSSF 356
Db 318 AFVGVKFRKMMFLTRLGRSD-----QRGPQRQPSRRRESSW 356
RESULT 13
CKR4 HUMAN
ID CKR4 HUMAN STANDARD; PRT; 360 AA.
AC P51679; Q9ULY7; Q9ULY7;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE C-C chemokine receptor type 4 (C-C CKR-4) (CCR-4) (CCR4)
DE (K5-5)
GN CCR4 OR CMKBR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Splicein;
RX MEDLINE=95370289; PubMed=7642634;
RA Power C.A., Meyer A., Nemeth K., Bacon K.B., Hoogewerf A.J.,
RA Proudfoot A.E.I., Wells T.N.C.;
RT "Molecular cloning and functional expression of a novel CC chemokine
RT receptor cDNA from a human basophilic cell line.";
RL J. Biol. Chem. 270:19495-19500(1995).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS VAL-130 AND SER-178.
RX MEDLINE=21040311; PubMed=11196669;
RA Kato H., Tsuchiya N., Izumi S., Miyamasu M., Nakajima T., Kawasaki H.,
RA Hirai K., Tokunaga K.;
RT "New variations of human CC-chemokine receptors CCR3 and CCR4.";
RL Genes Immun. 1:97-104(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RN Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.
RN [4]
RP FUNCTION.
RX MEDLINE=97313486; PubMed=9169480;
RA Imai T., Baba M., Nishimura M., Kakizaki M., Takagi S., Yoshie O.;
RT "The T cell-directed CC chemokine TARC is a highly specific
RT biological ligand for CC chemokine receptor 4.";
RL J. Biol. Chem. 272:15036-15042(1997).
RN [5]
RP FUNCTION.
RX MEDLINE=98104168; PubMed=9430724;
RA Imai T., Chantry D., Raport C.J., Wood C.L., Nishimura M., Godiska R.,
RA Yoshie O., Gray P.W.;
RT "Macrophage-derived chemokine is a functional ligand for the CC
RT chemokine receptor 4.";
RL J. Biol. Chem. 273:1764-1768(1998).
RN [6]
RP FUNCTION.
RX MEDLINE=99394604; PubMed=10466728;
RA Campbell J.J., Haraldsen G., Pan J., Rottman J., Qin S., Ponath P.,
RA Andrew D.P., Warnke R., Ruffing N., Kassam N., Wu L., Burcher E.C.;
RT "The chemokine receptor CCR4 in vascular recognition by cutaneous but
RT not intestinal memory T cells.";
RL Nature 400:776-780(1999).
RN [7]

RP FUNCTION, TISSUE SPECIFICITY, AND PHOSPHORYLATION.
RX MEDLINE=20219238; PubMed=10754297;
RA Ingberdingen M., Danaei B., Maghazachi A.A.;
RT "Human NK cells express CC chemokine receptors 4 and 8 and respond to
RT thymus and activation-regulated chemokine, macrophage-derived
RT chemokine, and I-309.";
RL J. Immunol. 164:4048-4054(2000).
CC -I- FUNCTION: High affinity receptor for the C-C type chemokines
CC TARC/SCYA17 and MDC/SCYA22. The activity of this receptor is
CC mediated by G(i) proteins which activate a phosphatidylinositol-
CC calcium second messenger system. Can function as a chemoattractant
CC homing receptor on circulating memory lymphocytes and as a
CC coreceptor for some primary HIV-2 isolates. In the CNS, could
CC mediate hippocampal-neuron survival.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: Predominantly expressed in the thymus, in
CC peripheral blood leukocytes, including T cells, mostly CD4+ cells,
CC and basophils, and in platelets; at lower levels, in the spleen
CC and in monocytes. Detected also in macrophages, IL-2-activated
CC natural killer cells and skin-homing memory T cells, mostly the
CC ones expressing the cutaneous lymphocyte antigen (CLA). Expressed
CC in brain microvascular and coronary artery endothelial cells.
CC -I- PTM: IN NATURAL KILLER CELLS, SCYA22 BINDING INDUCES
CC PHOSPHORYLATION ON YET UNDEFINED SER/THR RESIDUES, MOST PROBABLY
CC BY BETA-ADRENERGIC RECEPTOR KINASES 1 AND 2.
CC -I- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85740; CAA59743.1; -;
DR EMBL; AB023888; BAA8965.1; -;
DR EMBL; AB023889; BAA8966.1; -;
DR EMBL; AB023890; BAA8967.1; -;
DR EMBL; AB023891; BAA8968.1; -;
DR EMBL; AB023892; BAA8969.1; -;
DR EMBL; AV322539; AAP84352.1; -;
DR PIR; AS7160; AS7160.
DR Genew; HGNC:1605; CCR4.
DR MIM; 604936; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0004950; F:chemokine receptor activity; TAS.
DR GO; GO:0006935; P:chemotaxis; TAS.
DR GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0006954; P:inflammatory response; TAS.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW phosphorylation; Polymorphism.
FT DOMAIN 1 39 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 40 67 1 (POTENTIAL).
FT DOMAIN 68 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 98 2 (POTENTIAL).
FT DOMAIN 99 111 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 112 133 3 (POTENTIAL).
FT DOMAIN 134 150 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 151 175 4 (POTENTIAL).
FT DOMAIN 176 206 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 207 226 5 (POTENTIAL).
FT DOMAIN 227 242 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 243 267 6 (POTENTIAL).
FT DOMAIN 268 284 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 285 308 7 (POTENTIAL).
FT DOMAIN 309 360 CYTOPLASMIC (POTENTIAL).

```
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 110 187 BY SIMILARITY.
FT VARIANT 130 130 L -> V.
FT VARIANT 178 178 /FTID=VAR_010669.
FT VARIANT 178 178 C -> S.
FT SEQUENCE 360 AA; 41402 MW; 51EBE12AD1FAFABF CRC64;
Query Match 30.6%; Score 570.5; DB 1; Length 360;
Best Local Similarity 37.4%; Pred. No. 1.1e-30;
Matches 120; Conservative 63; Mismatches 125; Indels 13; Gaps 6;
QY 1 MEPSISVITSDYNEEVGSGD--YDSNKEPCFRDENVHFNRIPLTYFIPLTGIVGN 58
Db 1 MNPTDIAITLD--BSIYNNYLYESIPKPKTEGKIFAGELFLPLSLVVFVGLLGN 57
QY 59 GLVILVMGYOKLRSMTDKYRLHLSVADLLFVITLPFAVDAMADYVFGKFLCKAVHIY 118
Db 58 SVVVLVLFKYRLRSMTDVYLLNLAIADLLFVSLPFWGYAADQVFGLGCKMISWXY 117
QY 119 TNLVSSVLLIAFISLDRLYLAIHVAATNSQRPRLKLAEKAVYGVWIPALLLTIPDFAD 178
Db 118 LVGFYSGIFPVMLEMSDRYLALHVAFSLRARTLTGVITSLATWSAVFASLPLGFPST 177
QY 179 VSQGDISQGDRIYCDRLYP--DSLMMVVFQFHIMVGLIPGLVILSCYCIISKLSHS 236
Db 178 C-----YTERNHTY-CKTKYSLNSTTKVLSLSINILGLVPIGLMFCYSMIIRTLQHC 232
QY 237 KGHOKRKALKTTVILLIAPFACWLPYVVGISIDSFILLGVKOGCFESIVHKWISITRA 296
Db 233 KNEKKNKAVQMIYAVVVLFGFTPTNNIVLFELTVELEV-ODCTFERYLDAIQAET 291
QY 297 LAFPHCCNLPILYAFGAKPK 317
Db 292 LAFVHCCLNBIYVFLGCKPR 312
RESULT 14
IL8B RAT STANDARD; PRT; 359 AA.
AC P35407;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE High affinity interleukin-8 receptor B (IL-8R B) (CXCR-2) (GRO/MGSA
DE receptor).
DE Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1];
RN RATTUS FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Lung;
RX MEDLINE=97361587; PubMed=9218548;
RA Gobl A.E., Huang M.R., Wang S., Zhou Y., Oeberg K.;
RT "Molecular cloning and characterization of a cDNA encoding the rat
RT interleukin-8 receptor."
RL Biochim. Biophys. Acta 1326:171-177(1997).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RA Konishi K., Shibata F., Watanabe K., Tsurufuji S., Nakagawa H.,
RA Fujioaka M.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Lung, and Spleen;
RX MEDLINE=97115810; PubMed=8955112;
RA Dunstan C.-A.N., Salafra M.N., Adhikari S., Xia Y., Feng L.,
RA Harrison J.K.;
RT "Identification of two rat genes orthologous to the human
```

```
interleukin-8 receptors."
J. Biol. Chem. 271:32770-32776(1996).
-!- FUNCTION: Receptor to interleukin-8, which is a powerful
neutrophils chemotactic factor. Binding of IL-8 to the receptor
causes activation of neutrophils. This response is mediated via a
G-protein that activate a phosphatidylinositol-calcium second
messenger system. This receptor binds to IL-8 with a high affinity
and to GRO/MGSA and NAP-2 also with a high affinity.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X77797; CAA54824.1; -
EMBL; D63584; BAA09797.1; -
EMBL; U70988; AAC52961.1; -
InterPro; IPR000276; GPCR_Rhodopsn.
Pfam; PF00001; 7tm.1; 1.
PRINTS; PR00237; GPCRHOPOPSN
PROSITE; PS00237; G-PROTEIN RECP F1.1; 1.
PROSITE; PS00262; G-PROTEIN RECP F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Chemotaxis.
FT DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 48 74 1 (POTENTIAL).
FT DOMAIN 75 83 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 84 104 2 (POTENTIAL).
FT DOMAIN 105 119 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 120 141 3 (POTENTIAL).
FT DOMAIN 142 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 182 4 (POTENTIAL).
FT DOMAIN 183 207 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 208 230 5 (POTENTIAL).
FT DOMAIN 231 250 6 (POTENTIAL).
FT TRANSMEM 251 272 7 (POTENTIAL).
FT DOMAIN 273 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 314 7 (POTENTIAL).
FT DOMAIN 315 359 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 118 195 BY SIMILARITY.
SQ SEQUENCE 359 AA; 40532 MW; EF547326A074ABDD CRC64;
Query Match 30.1%; Score 562; DB 1; Length 359;
Best Local Similarity 35.0%; Pred. No. 4e-30;
Matches 130; Conservative 67; Mismatches 126; Indels 48; Gaps 10;
QY 12 DNYS-EEVSGSDYD-----SNKEPCFRDENVHFNRIPLTYFIPLTGIVGN 58
Db 7 DNFSLEDFSGDIDSYNSVSDPPFTLSDAAPC-PSANLDINRYAVVYVLTLLSLVGN 65
QY 59 GLVILVMGYOKLRSMTDKYRLHLSVADLLFVITLPFAVDAMADYVFGKFLCKAVHIY 118
Db 66 SLVNLVILYNRESTCVTDVLLNLAIADLPFALTLPVNAASKVNGWIFGSLCKVFSFLQ 125
QY 119 TNLVSSVLLIAFISLDRLYLAIHVAATNSQRPRLKLAEKAVYGVWIPALLLTIPDFAD 178
Db 126 EITFYSSVLLIACISMDRYLAIVHATSTLIQKRLH-V-KVCITWFLSLVLSIFIL-- 182
QY 179 VSQGDISQGDRIYCDRLYP-----SLMMVVFQFHIMVGLIPGLVILSCYCIISKL 233
Db 183 --RTTKANPSTVVC---YENIGNNTSKRWVRLPLPQTYGFLPLLMFLFCYFTIRL 237
QY 234 SHSKGHOKRKALKTTVILLIAPFACWLPYVVGISIDSFILLGVKOGCFESIVHKWISI 293
```

Db 238 FXAHMGQKRAMRVJFAVVLVFLLCWLPYNVILFTDTLMRTKLKETCERONEINKALEA 297
Qy 294 TEALAFHCCCLNPILYAFGLAKFKSAQHALNSMRGSSLSKLSK-----GREGGH 344
Db 298 TEILGFLHSLNPIIYAFIGQKFR-----HGL-----LKIMANYGLVSKFLAKEGRP 345
Qy 345 SSVSTESSESS 355
Db 346 SPVGSSANTS 356

RESULT 15
IL8A RABIT
ID IL8A RABIT STANDARD; PRT; 355 AA.
AC P21109;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE High affinity interleukin-8 receptor A (IL-8R A) (CXCR-1).
GN IL8RA OR CXCR1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91378994; PubMed=1898400;
RA Beckmann M.P., Munger W.E., Kozlosky C., Varden Bos T., Price V.,
RA Lyman S., Gerard N.P., Gerard C., Czerretti D.P.;
RT "Molecular characterization of the interleukin-8 receptor.";
RL Biochem. Biophys. Res. Commun. 179:784-789(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Albino; TISSUE=Neutrophils;
RX MEDLINE=91056034; PubMed=1700779;
RA Thomas K.M., Pyun H.Y., Navarro J.;
RT "Molecular cloning of the fMet-Leu-Phe receptor from neutrophils.";
RL J. Biol. Chem. 265:20061-20064(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Neutrophils;
RX MEDLINE=92148149; PubMed=1737938;
RA Lee J., Kuang W.-J., Rice G.C., Wood W.I.;
RT "Characterization of complementary DNA clones encoding the rabbit
IL-8 receptor.";
RL J. Immunol. 148:1261-1264(1992).
CC -!- FUNCTION: Receptor to interleukin-8, which is a powerful
neutrophils chemotactic factor. Binding of IL-8 to the receptor
causes activation of neutrophils. This response is mediated via a
G-protein that activate a phosphatidylinositol-calcium second
messenger system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Neutrophils.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -!- CAUTION: Was originally (Ref.2) thought to be the receptor for
fMet-Leu-Phe (N-formyl peptide receptor).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC ENBL; M74240; AAA31375.1; -.
DR ENBL; M58021; AAA31377.1; -.
DR ENBL; M82873; AAA31376.1; -.
DR PIR; A23669; A23669.
DR PIR; J01231; J01231.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; P800237; G-PROTEIN RECEPTOR; 1.
DR PROSITE; P800262; G-PROTEIN RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Chemotaxis.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 67 1 (POTENTIAL).
FT DOMAIN 68 73 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 74 92 2 (POTENTIAL).
FT DOMAIN 93 114 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 115 138 3 (POTENTIAL).
FT DOMAIN 139 159 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 160 184 4 (POTENTIAL).
FT DOMAIN 185 204 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 205 232 5 (POTENTIAL).
FT DOMAIN 233 247 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 248 270 6 (POTENTIAL).
FT DOMAIN 271 290 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 291 313 7 (POTENTIAL).
FT DOMAIN 314 355 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 7 21 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 21 121 BY SIMILARITY.
FT DISULFID 115 192 DILFALTPWIAVSKWKWIFG -> PAFCPDPAHYLGRDQGRIDFR (IN REF. 2).
FT CONFLICT 146 147 HA -> QS (IN REF. 2).
FT CONFLICT 204 204 R -> C (IN REF. 2).
FT CONFLICT 287 288 DI -> EL (IN REF. 2).
SQ SEQUENCE 355 AA; 40622 MW; EFB49ACB9D1E0F21 CRC64;

Query Match 29.5%; Score 551.5; DB 1; Length 355;
Best Local Similarity 35.0%; Pred. No. 1.9e-29;
Matches 121; Conservative 72; Mismatches 142; Indels 11; Gaps 6;

Qy 12 DNYSEEVSGDYDSNKEPCFRDENVHFNRIPTIYFIIFLTGIVGNGLVILVMGYQKKL 71
Db 17 DEANATGMPPEKDYSPCLVVTQT-LNKYVVVVIYALVFLSLGNSLVMLVILYSRN 75
Qy 72 RSMYDXYRLHLSVADLLFVITLPPFAVDAMADWYFKFLCKAVHIIYTNLYSSVLLAF 131
Db 76 RSVTDVYLLNLAWADLLFALTPWIAVSKWKWIFGTPLCKVSVLKYFNFGILLAC 135
Qy 132 ISLDYLAIVHATNSQRPRLKLAEKAVYGVWIPALLTIPDFIADVSQGDISQDDRY 191
Db 136 ISVDYLAIVHATRTLTQKRLV-KFICIGIUALSLILSPFLFRQV----FSPNNSP 190
Qy 192 IC--DRLYPDSLMMVVVFQFHIMVGLILPGVILSCYCIISKLKSHSKGKRAKTKTV 249
Db 191 VCYEDLGHNTAKRWVLRILPHTFGFILLVLMFCYGTLTTLFOAHMGQKRAMRVIF 250
Qy 250 ILILAFACWLPYVVGISIDSPILLGVIKQCDPESIVHKWISITEALAFFHCCCLNPILY 309
Db 251 AVVLIIFLLCWLPLYNEVLLADTLMTHTVIGETCORNDIDRALDADTEILGFLHSLNPIY 310
Qy 310 AFLGAKFKSSAQAHALNSMRG-SSLKILSKGKRGHSSVSTESSESS 354
Db 311 AFIQGNFRNGFLKWL--AAGLISKELTRHRTVSTSSSTNVPSN 354

Search completed: May 17, 2004, 22:16:44
Job time : 28 secs

OM protein - protein search, using sw model

Run on: May 17, 2004, 21:20:57 ; Search time 128 Seconds
(without alignments)
884.930 Million cell updates/sec

Title: US-09-367-052-2
Perfect score: 1867
Sequence: 1 MPEISVSYTSDNYSEEVGS.....KGGHSSVSTSESSSPHSS 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_protist:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1758.5	94.2	349	11 Q8VD47	Q8vd47 rattus norv
2	1692.5	90.7	352	6 Q7YS92	Q7ys92 tupaiia bela
3	1691.5	90.6	352	6 Q8HZU1	Q8hzu1 callithrix
4	1689.5	90.5	352	6 Q8HZU0	Q8hzu0 saimiri sci
5	1679.5	90.0	352	6 Q77488	Q77488 cercopithec
6	1678.5	89.9	347	6 Q9MZM9	Q9mzm9 ateles pani
7	1677.5	89.9	347	6 Q9MZP7	Q9mzp7 presbytis s
8	1676.5	89.8	347	6 Q9MZP8	Q9mzp8 colobus pol
9	1676.5	89.8	347	6 Q9MZQ0	Q9mzq0 pygathrix n
10	1676.5	89.8	347	6 Q9MZM5	Q9mzm5 pithecia pi
11	1676.5	89.8	347	6 Q9MZP0	Q9mzp0 macaca assa
12	1676.5	89.8	347	6 Q9MZQ2	Q9mzq2 pygathrix r
13	1676.5	89.8	347	6 Q9MZP4	Q9mzp4 presbytis f
14	1676.5	89.8	347	6 Q9MZP9	Q9mzp9 nasalis lar
15	1676.5	89.8	347	6 Q9MZP5	Q9mzp5 presbytis p
16	1676.5	89.8	347	6 Q9MZQ3	Q9mzq3 pygathrix a

17	1676.5	89.8	347	6 Q9MZN8	Q9mzn8 macaca thib
18	1676.5	89.8	347	6 Q9MZN9	Q9mzn9 macaca arct
19	1676.5	89.8	347	6 Q9MZN7	Q9mzn7 macaca neme
20	1676.5	89.8	347	6 Q9MZN1	Q9mzn1 pygathrix b
21	1676.5	89.8	347	6 Q9MZN1	Q9mzn1 pygathrix b
22	1676.5	89.8	347	6 Q9MZN1	Q9mzn1 cercopithec
23	1676.5	89.8	347	6 Q9MZN1	Q9mzn1 eulemur mac
24	1676.5	89.7	347	6 Q9MZN0	Q9mzn0 alouatta se
25	1676.5	89.6	347	6 Q9MZN2	Q9mzn2 mandrillus
26	1676.5	89.6	347	6 Q9MZN2	Q9mzn2 macaca fasc
27	1676.5	89.5	347	6 Q9MZN4	Q9mzn4 hylobates l
28	1676.5	89.5	347	6 Q9MZN6	Q9mzn6 hylobates c
29	1676.5	89.5	347	6 Q9MZN5	Q9mzn5 hylobates h
30	1676.5	89.5	347	6 Q9MZN1	Q9mzn1 gorilla gor
31	1676.5	89.5	347	6 Q9MZN2	Q9mzn2 pongo pygma
32	1676.5	89.5	347	6 Q9MZN2	Q9mzn2 homo sapien
33	1668.5	89.4	347	6 Q9MZN6	Q9mzn6 callicebus
34	1668.5	89.4	347	6 Q9MZN3	Q9mzn3 presbytis e
35	1668.5	89.4	347	6 Q9MZN3	Q9mzn3 presbytis j
36	1665.5	89.2	347	6 Q9MZN3	Q9mzn3 hylobates s
37	1663.5	89.1	347	6 Q9MZN8	Q9mzn8 callithrix
38	1661.5	89.0	347	6 Q9MZN1	Q9mzn1 mandrillus
39	1644.5	88.1	347	6 Q9MZN0	Q9mzn0 perodicticu
40	1638	87.7	346	6 Q9MZN2	Q9mzn2 nycticebus
41	1638	87.7	346	6 Q9MZN3	Q9mzn3 nycticebus
42	1638	87.7	346	6 Q9MZN4	Q9mzn4 nycticebus
43	1613	86.4	332	11 Q62973	Q62973 rattus norv
44	1611.5	86.3	334	6 Q9N1P2	Q9n1p2 saguinus oe
45	1606.5	86.0	334	6 Q9N130	Q9n130 saimiri bol

ALIGNMENTS

RESULT 1

Q8VD47	PRELIMINARY;	PRT;	349 AA.
ID	Q8VD47		
AC	Q8VD47;		
DT	01-MAR-2002 (Tremblrel. 20, Created)		
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)		
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)		
DE	Chemokine receptor CXCR4.		
GN	CXCR4.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Holtzman; TISSUE=Brain;		
RA	Simen A.A., Miller R.J.;		
RT	"Chemokine regulation of neuronal signaling and gpl20 neurotoxicity."		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
DR	EMBL; AF452185; AAL47855.1; -		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.		
DR	GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.		
DR	InterPro; IPR000276; GPCR_Rhodpsn.		
DR	InterPro; IPR003006; IG_MHC.		
DR	Pfam; PF00001; 7tm.1; 1		
DR	PRINTS; PR00237; GPCR_RHODPSN.		
DR	PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.		
DR	PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.		
DR	PROSITE; PS00290; IG_MHC; 1.		
KW	G-protein coupled receptor; Receptor; Transmembrane.		
SQ	SEQUENCE 349 AA; 39429 MW; 09D19860D3D2CB8A CRC64;		

Query Match 94.2%; Score 1758.5; DB 11; Length 349;
Best Local Similarity 96.0%; Pred. No. 7.8e-150;
Matches 340; Conservative 4; Mismatches 5; Indels 5; Gaps 1;


```

DR GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO:00007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR003006; Ig MHC.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
DR PROSITE; PS00290; IG MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON TER
SQ SEQUENCE 347 AA; 39317 MW; B84037A04490A1D1 CRC64;

Query Match      89.9%; Score 1678.5; DB 6; Length 347;
Best Local Similarity 90.6%; Pred. No. 1.2e-142;
Matches 319; Conservative 18; Mismatches 10; Indels 5; Gaps 1;

QY 8 IYTSNYSVEVSGDYSNKEPCFRDENVHFNRIPLPTIYFIIFLTGIVGNGLVILVMGY 67
Db 1 IYTSNYSVEVSGDYSNKEPCFRDENVHFNRIPLPTIYFIIFLTGIVGNGLVILVMGY 60
QY 68 QKLRSMTDKYLHLHSVADLLFVITLPFWADVADAWYFGKFLCKAVHIIYTNLYSSVL 127
Db 61 QKLRSMTDKYLHLHSVADLLFVITLPFWADVADAWYFGKFLCKAVHIIYTNLYSSVL 120
QY 128 ILAFISLDRLYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIADVSQGDISQ 187
Db 121 ILAFISLDRLYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIADVSQGDISQ 175
QY 188 DRYICDRILYPSLMMVVPQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 247
Db 176 EDRIYCDRYPNDLMMVVPQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 235
QY 248 TVILLIAPFACWLPYVYGISIDSFILLGVIKGCDPFESIVHKWISITTEALAFHCCLNPI 307
Db 236 TVILLIAPFACWLPYVYGISIDSFILLGVIKGCDPFESIVHKWISITTEALAFHCCLNPI 295
QY 308 LYAFLGAKPKSAQAHALNSMRGSSSLKILSKRGHSHSVSTESSSFHSS 359
Db 296 LYAFLGAKPKTSAQHALTSVSRGSSSLKILSKRGHSHSVSTESSSFHSS 347

RESULT 7
Q9MZP7 PRELIMINARY; PRT; 347 AA.
AC Q9MZP7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Presbytis senex.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Presbytis.
OX NCBI_TaxID=66055;
RN [1]_TaxID=66055;
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF172216; AAF89336.1; -.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0004872; F:receptor activity; IEA.
DR GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR003006; GPCR Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.

```

```

DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
DR PROSITE; PS00290; IG MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON TER
SQ SEQUENCE 347 AA; 39191 MW; 2F69F0CA9CBC3878 CRC64;

Query Match      89.9%; Score 1677.5; DB 6; Length 347;
Best Local Similarity 90.3%; Pred. No. 1.5e-142;
Matches 318; Conservative 17; Mismatches 12; Indels 5; Gaps 1;

QY 8 IYTSNYSVEVSGDYSNKEPCFRDENVHFNRIPLPTIYFIIFLTGIVGNGLVILVMGY 67
Db 1 IYTSNYSVEVSGDYSNKEPCFRDENVHFNRIPLPTIYFIIFLTGIVGNGLVILVMGY 60
QY 68 QKLRSMTDKYLHLHSVADLLFVITLPFWADVADAWYFGKFLCKAVHIIYTNLYSSVL 127
Db 61 QKLRSMTDKYLHLHSVADLLFVITLPFWADVADAWYFGKFLCKAVHIIYTNLYSSVL 120
QY 128 ILAFISLDRLYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIADVSQGDISQ 187
Db 121 ILAFISLDRLYLAIVHATNSQRPKLLAEKAVVGVWIPALLLTIPDFIADVSQGDISQ 175
QY 188 DRYICDRILYPSLMMVVPQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 247
Db 176 DRYICDRYPNDLMMVVPQFQHMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 235
QY 248 TVILLIAPFACWLPYVYGISIDSFILLGVIKGCDPFESIVHKWISITTEALAFHCCLNPI 307
Db 236 TVILLIAPFACWLPYVYGISIDSFILLGVIKGCDPFESIVHKWISITTEALAFHCCLNPI 295
QY 308 LYAFLGAKPKSAQAHALNSMRGSSSLKILSKRGHSHSVSTESSSFHSS 359
Db 296 LYAFLGAKPKTSAQHALTSVSRGSSSLKILSKRGHSHSVSTESSSFHSS 347

RESULT 8
Q9MZP8 PRELIMINARY; PRT; 347 AA.
AC Q9MZP8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Colobus polykomos.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Colobus.
OX NCBI_TaxID=9572;
RN [1]_TaxID=9572;
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF172215; AAF89335.1; -.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0004872; F:receptor activity; IEA.
DR GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR003006; Ig MHC.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
DR PROSITE; PS00290; IG MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON TER
SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;

```

[illegible]

QY 128 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 187
Db 121 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 175
QY 188 DRYICDRYLPDLWVWVQFQHIMVGLILPGIVILSCYCIISKLSSHKGQKRAKLT 247
Db 176 DRYICDRYLPDLWVWVQFQHIMVGLILPGIVILSCYCIISKLSSHKGQKRAKLT 235
QY 248 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 359
Db 236 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 347
QY 308 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 359
Db 296 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 347

RESULT 11

Q9MZP0 PRELIMINARY; PRT; 347 AA.
ID Q9MZP0
AC Q9MZP0
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Macaca assamensis (Assam's macaque) (Assam's monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9551;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF172223; AAF89343.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;

Query Match 89.8%; Score 1676.5; DB 6; Length 347;
Best Local Similarity 90.6%; Pred. No. 1.8e-142;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

QY 8 IYTSNYSSEVSGDYDSNKEPCFRDENVHFRIFLPTIYFIIFLTGIVGNGLVILWGY 67
Db 1 IYTSNYSSEVSGDYDSNKEPCFRDENVHFRIFLPTIYFIIFLTGIVGNGLVILWGY 60
QY 68 QKLRSMTDKYLHLVADLLFVITLFFWADVADWYFGKFLCAVHIITVNLVSSVL 127
Db 61 QKLRSMTDKYLHLVADLLFVITLFFWADVADWYFGKFLCAVHIITVNLVSSVL 120
QY 128 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 187
Db 121 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 175
QY 188 DRYICDRYLPDLWVWVQFQHIMVGLILPGIVILSCYCIISKLSSHKGQKRAKLT 247
Db 236 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 359
QY 296 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 347

Db 176 DRYICDRYLPDLWVWVQFQHIMVGLILPGIVILSCYCIISKLSSHKGQKRAKLT 235
QY 248 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 359
Db 236 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 347
QY 308 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 359
Db 296 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 347

RESULT 12

Q9MZQ2 PRELIMINARY; PRT; 347 AA.
ID Q9MZQ2
AC Q9MZQ2
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Pygathrix roxellana (golden snub-nosed monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.
OX NCBI_TaxID=61622;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF172211; AAF89331.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;

Query Match 89.8%; Score 1676.5; DB 6; Length 347;
Best Local Similarity 90.6%; Pred. No. 1.8e-142;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

QY 8 IYTSNYSSEVSGDYDSNKEPCFRDENVHFRIFLPTIYFIIFLTGIVGNGLVILWGY 67
Db 1 IYTSNYSSEVSGDYDSNKEPCFRDENVHFRIFLPTIYFIIFLTGIVGNGLVILWGY 60
QY 68 QKLRSMTDKYLHLVADLLFVITLFFWADVADWYFGKFLCAVHIITVNLVSSVL 127
Db 61 QKLRSMTDKYLHLVADLLFVITLFFWADVADWYFGKFLCAVHIITVNLVSSVL 120
QY 128 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 187
Db 121 ILAFISLDRLYLAIATVHATNSQRPRKLAEKAVVGVWIPALLLTTPDRIADVSQGDISQ 175
QY 188 DRYICDRYLPDLWVWVQFQHIMVGLILPGIVILSCYCIISKLSSHKGQKRAKLT 247
Db 236 TVILLILAFACWLPYVYVIGSIDSFILLGKILSKRGHSHSVSTESSSPHSS 359
QY 296 LYAFILGAKFKSAQAHLNMSRGSLSKILSKRGHSHSVSTESSSPHSS 347

QY 308 LYAFGLGAKFKSSAQHALNSMRGSSKILSKGRGGHSSVSTESSESSPHSS 359
Db 296 LYAFGLGAKFKTSQAQHALTSVSRGSSKILSKGRGGHSSVSTESSESSPHSS 347

RESULT 13

Q9MZP4 ID Q9MZP4 PRELIMINARY; PRT; 347 AA.
AC Q9MZP4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Presbytis francoisi.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Presbytis.
OX NCBI_TaxID=33549;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF172219; AAF89339.1;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GFCRRHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON TER 1
SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;

Query Match 89.8%; Score 1676.5; DB 6; Length 347;
Best Local Similarity 90.6%; Pred. No. 1.8e-142;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

QY 8 IYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTYIFLTGIVGNGLVILVMGY 67
Db 1 IYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTYIFLTGIVGNGLVILVMGY 60
QY 68 QKXLRMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIYTNLYSSVL 127
Db 61 QKXLRMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIYTNLYSSVL 120
QY 128 ILAFISLDRLYLAI VHAATNSORPKLAEKAVYGVWIPALLTIPDFIFADVSQGDISQ 187
Db 121 ILAFISLDRLYLAI VHAATNSORPKLAEKAVYGVWIPALLTIPDFIFASV-----SEA 175
QY 188 DDRIYCDRLYPDSLWVWVFOQHIMVGLILPGVILSCYCIILSKLSHGKHQKALKT 247
Db 176 DDRIYCDRFYNDLWVWVFOQHIMVGLILPGVILSCYCIILSKLSHGKHQKALKT 235
QY 248 TVIILILAFFACWLPYYVIGISIDSFILLGVTKQCDPESIVHKWISITTEALAFHCCLNPI 307
Db 236 TVIILILAFFACWLPYYVIGISIDSFILLGVTKQCEFTVHKWISITTEALAFHCCLNPI 295
QY 308 LYAFGLGAKFKSSAQHALNSMRGSSKILSKGRGGHSSVSTESSESSPHSS 359
Db 296 LYAFGLGAKFKTSQAQHALTSVSRGSSKILSKGRGGHSSVSTESSESSPHSS 347

RESULT 15

Q9MZP5 ID Q9MZP5 PRELIMINARY; PRT; 347 AA.
AC Q9MZP5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

RESULT 14

Q9MZP9 ID Q9MZP9 PRELIMINARY; PRT; 347 AA.
AC Q9MZP9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Chemokine receptor CXCR4 (Fragment).
GN CXCR4.
OS Nasalis larvatus (Proboscis monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Nasalis.
OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
RT "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in Primates."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF172214; AAF89334.1;
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GFCRRHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
FT NON TER 1
SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;

Query Match 89.8%; Score 1676.5; DB 6; Length 347;
Best Local Similarity 90.6%; Pred. No. 1.8e-142;
Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;

QY 8 IYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTYIFLTGIVGNGLVILVMGY 67
Db 1 IYTSNYSVEVSGDYDSNKEPCFRDENVHFNRIPLTYIFLTGIVGNGLVILVMGY 60
QY 68 QKXLRMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIYTNLYSSVL 127
Db 61 QKXLRMTDKYRLHLSVADLLFVITLFPWAVDAMADWYFGKFLCKAVHIYTNLYSSVL 120
QY 128 ILAFISLDRLYLAI VHAATNSORPKLAEKAVYGVWIPALLTIPDFIFADVSQGDISQ 187
Db 121 ILAFISLDRLYLAI VHAATNSORPKLAEKAVYGVWIPALLTIPDFIFASV-----SEA 175
QY 188 DDRIYCDRLYPDSLWVWVFOQHIMVGLILPGVILSCYCIILSKLSHGKHQKALKT 247
Db 176 DDRIYCDRFYNDLWVWVFOQHIMVGLILPGVILSCYCIILSKLSHGKHQKALKT 235
QY 248 TVIILILAFFACWLPYYVIGISIDSFILLGVTKQCDPESIVHKWISITTEALAFHCCLNPI 307
Db 236 TVIILILAFFACWLPYYVIGISIDSFILLGVTKQCEFTVHKWISITTEALAFHCCLNPI 295
QY 308 LYAFGLGAKFKSSAQHALNSMRGSSKILSKGRGGHSSVSTESSESSPHSS 359
Db 296 LYAFGLGAKFKTSQAQHALTSVSRGSSKILSKGRGGHSSVSTESSESSPHSS 347

01-JUN-2003 (Tremblrel. 24, Last annotation update)
 Chemokine receptor CXCR4 (Fragment).
 GN CXCR4.
 OS Presbytis phayrei.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Presbytis.
 OC Presbytis.
 OX NCBI_TaxID=66054;
 [1]
 RN SEQUENCE FROM N.A.
 RP Zheng X.-Z., Zhang Y.-P., Geng Z.-C.;
 RA "Molecular Evolution of the Chemokine Receptor CXCR4 Gene Sequences in
 RT Primates.",
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF172218; AAF89338.1; ..
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 347 AA; 39205 MW; E96B66136541515C CRC64;
 Query Match 89.8%; Score 1676.5; DB 6; Length 347;
 Best Local Similarity 90.6%; Pred. No. 1.8e-142;
 Matches 319; Conservative 16; Mismatches 12; Indels 5; Gaps 1;
 QY 8 IYTSNYSEVGSGDYDSNKEPCFRDENVHFNRIELPTIYFIIFLTGIVGNGLVILVMGY 67
 DB 1 IYTSNYTEEMSGDYDSIKEPCFRENAHFNRIELPTIYFIIFLTGIVGNGLVILVMGY 60
 QY 68 QXKLSMTDKYRLHLSVADLLVITLPFWADVADADWYFGKFLCKAVHIIYTNLYSSVL 127
 DB 61 QXKLSMTDKYRLHLSVADLLVITLPFWADVADADWYFGKFLCKAVHIIYTNLYSSVL 120
 QY 128 ILAFISLDRLAIVHATNSORPKLLAEKAVYVGVWIPALLITIPDFIFADVSGDISQG 187
 DB 121 ILAFISLDRLAIVHATNSQKPKLLAEKAVYVGVWIPALLITIPDFIFASV-----SEA 175
 QY 188 DRYICDRLYPDSLMWVFPQHIMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 247
 DB 176 DRYICDRLYPNDLMWVFPQHIMVGLILPGIVILSCYCIISKLSHSGHQKRAKLT 235
 QY 248 TVILLIARFACWLPYVVGISIDSPILLGVIKGCDPESIVHKWISITTEALAFHCCLNPI 307
 DB 236 TVILLIARFACWLPYVVGISIDSPILLGVIKGCDPESIVHKWISITTEALAFHCCLNPI 295
 QY 308 LYAFLGAKFKSAQAHALNSMRGSSLKILSKGRGHSVSTESSESSSFHSS 359
 DB 296 LYAFLGAKFKSAQAHALTSVRGSSLKILSKGRGHSVSTESSESSSFHSS 347

Search completed: May 17, 2004, 22:19:03
 Job time : 131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 17, 2004, 04:56:59 ; Search time 2770.88 Seconds
(without alignments)
10714.982 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccaatcctaatacgaactact.....gacatcagtcagggggatga 685

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl: *
- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	685	100.0	1877	10	D87747	D87747 Mus musculus
2	636	92.8	1817	10	BC031665	BC031665 Mus musculus
3	631.2	92.1	1575	10	MMLESTRPT	X99582 M.musculus
4	618.6	90.3	1809	10	AB000803	AB000803 Mouse mRN
5	595	86.9	1223	10	MMU59760	U59760 Mus musculus
6	548.4	80.1	1180	10	MMLCR12	Z80112 Mus musculus
7	545.8	79.7	1422	10	MMLCR13	Z80113 Mus musculus
8	545.8	79.7	3366	10	MMU65380	U65380 Mus musculus
9	541	79.0	3770	10	MMLESTRGN	X99581 M.musculus
10	541	79.0	175370	2	AC145861	AC145861 Pan trogl
11	541	79.0	185215	2	AC147556	AC147556 Mus muscu
12	495.8	72.4	528	10	MMLCR11	Z80111 Mus musculus
13	473.4	69.1	190149	2	AC122097	AC122097 Rattus no
14	472.4	69.0	1050	10	AF452185	AF452185 Rattus no
15	469.2	68.5	1050	10	RN090610	U90610 Rattus norv
16	452.8	66.1	1126	9	AF019378	AF019378 Cercopith
17	446.8	65.2	1667	9	AK129916	AK129916 Homo sapi
18	446.8	65.2	1679	6	AX287116	AX287116 Sequence
19	446.8	65.2	1679	6	AX548790	AX548790 Sequence
20	446.8	65.2	1679	6	AX705078	AX705078 Sequence
21	445.8	65.1	1664	6	E37853	E37853 Anti-HIV ag
22	445.8	65.1	1664	6	AR222504	AR222504 Sequence
23	445.8	65.1	1664	6	AR380668	AR380668 Sequence
24	445.8	65.1	1664	9	HUMH89	DI0924 Homo sapien
25	445.8	65.1	1664	11	G28514	G28514 SHGC-31461
26	444.8	64.9	1637	9	HUMSTSR	M99293 Homo sapien
27	441.6	64.5	1645	9	HSNBYRLA	X71635 H.sapiens m
28	440.4	64.3	1059	9	AF031089	AF031089 Papio ham
29	439.2	64.1	1737	6	AR015970	AR015970 Sequence
30	439.2	64.1	1737	6	AR060748	AR060748 Sequence
31	439.2	64.1	1737	6	AR070433	AR070433 Sequence
32	439.2	64.1	1737	6	AR103430	AR103430 Sequence
33	439.2	64.1	1737	6	I13753	I13753 Sequence 4
34	439.2	64.1	1737	6	I24455	I24455 Sequence 3
35	438.8	64.1	1068	9	AB015943	AB015943 Chloroceb
36	437.6	63.9	1670	6	AX409496	AX409496 Sequence
37	437.6	63.9	1670	6	AX663641	AX663641 Sequence
38	437.6	63.9	1670	9	HUMGPCR	IO6797 Human (clon
39	437.2	63.8	1662	9	BC020968	BC020968 Homo sapi
40	435.6	63.6	1078	9	AF291672	AF291672 Macaca fa
41	434	63.4	1059	9	AF452613	AF452613 Saimiri s
42	434	63.4	1059	9	MMU73740	U73740 Macaca mula
43	434	63.4	1078	9	D86579	D86579 Macaca fasc
44	433.8	63.3	1044	9	AF172212	AF172212 Pygathrix
45	433.8	63.3	1044	9	AF172238	AF172238 Pithecia

ALIGNMENTS

RESULT 1
D87747
LOCUS D87747 1877 bp mRNA linear ROD 07-FEB-1999
DEFINITION Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION D87747
VERSION D87747.1 GI:1772445
KEYWORDS murine CXCR-4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (sites)
REFERENCE
Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
Yoshie,O., Matsushima,K., Yoshida,N., Springer,I.A. and
Kishimoto,T.

TITLE Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)

MEDLINE 97121456

PUBMED 8962122

REFERENCE 2 (bases 1 to 1877)

AUTHORS Nagasawa, T.

TITLE Molecular cloning of murine PBSF/SDF-1 receptor

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1877)

AUTHORS Nagasawa, T.

TITLE Direct Submission

JOURNAL Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka Medical Center, Department of Immunology; Murodocho 840, Isumi, Osaka 590-02, Japan (E-mail: immunol@osk.threewebnet.or.jp, Tel:0725-56-1220, Fax:0725-57-3021)

FEATURES Location/Qualifiers

1..1877

organism="Mus musculus"

mol_type="mRNA"

db_xref="taxon:10090"

cell_line="DW34"

cell_type="pre-B cell"

tissue_type="bone marrow"

1..1877

gene="murine CXCR-4"

120..1199

gene="murine CXCR-4"

function="PBSF/SDF-1 receptor"

codon_start=1

evidence=experimental

product="murine CXCR-4"

protein_id="BAA13451.1"

db_xref="GI:1772446"

translation="MEPISVSIYTSNDYSEVSGSDVSNKEPCFRDENVHFNRIPLTIYFIITGVNGVLVIMVGYQKLSMTDKYRLHSLVADLFIITLFPWADVADMA DWYFKFLKCAVHIITVYSSVILAFISLDYLAIVHATNSQRPKLLAEKAVV GVWIPALLITPDIYFADVSQDTSQDDRVICDLVPSLSLWVYVFOHITWGLIPL GVILSCYCIILISLHSHQKQKALATWILLILAFACWLPYVVISIDISFLLGV IKQGDPSFIVHKSISIEALAFPHCCLNPLIYAFGLAKFKSSAQHALNSRGSLSK ILKSGKRGHSGSVSTESSESSSFHS"

ORIGIN

Query Match 100.0%; Score 685; DB 10; Length 1877;

Best Local Similarity 100.0%; Pred. No. 5.8e-170;

Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCTAATAGCATCTACTATAGGGCTCGAGCGGCCCGCGGCGAGTGCAGTAGCAG 60

DB 1 CCATCTAATAGCATCTACTATAGGGCTCGAGCGGCCCGCGGCGAGTGCAGTAGCAG 60

QY 61 TGACCTCTGAGCGCTTGGTCTCGGTACCCACCGGCTGTAGCGAGTGTGCCA 120

DB 61 TGACCTCTGAGCGCTTGGTCTCGGTACCCACCGGCTGTAGCGAGTGTGCCA 120

QY 121 TGGAAACCGATCAGTGTGAGTATATACATCTTGTGATACTCTGAAAGTGGGGTGTG 180

DB 121 TGGAAACCGATCAGTGTGAGTATATACATCTTGTGATACTCTGAAAGTGGGGTGTG 180

QY 181 GAGATATGATCCCAAGAGAACCTGTTCCGGATGAAACGTCATTCATAGGA 240

DB 181 GAGATATGATCCCAAGAGAACCTGTTCCGGATGAAACGTCATTCATAGGA 240

QY 241 TCTTCTGCCACCATCTACTTTCATCATCTTCTTGTGATGATGCGGCAATGG 300

DB 241 TCTTCTGCCACCATCTACTTTCATCATCTTCTTGTGATGATGCGGCAATGG 300

QY 301 TGATCTGTGTGATGGTGTACGAGAGAGCTAAGAGCATGACGGAACGATCGGCTGC 360

DB 301 TGATCTGTGTGATGGTGTACGAGAGAGCTAAGAGCATGACGGAACGATCGGCTGC 360

QY 361 ACCTGTCACTGGCTGACCTCTCTTTTGTGATCACAACCTCCCTTGGGAGTTGATGCCA 420

DB 361 ACCTGTCACTGGCTGACCTCTCTTTTGTGATCACAACCTCCCTTGGGAGTTGATGCCA 420

QY 421 TGCGTCACTGGTACTTTGGGAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 480

DB 421 TGCGTCACTGGTACTTTGGGAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 480

QY 481 ACCTTACAGCAGCGTTCTCATCTGGCTTTCATCAGCTGACCGGTACCTCGCATTTG 540

DB 481 ACCTTACAGCAGCGTTCTCATCTGGCTTTCATCAGCTGACCGGTACCTCGCATTTG 540

QY 541 TCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAAGGAGTCTATGTGG 600

DB 541 TCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGCTGAAAAGGAGTCTATGTGG 600

QY 601 GCGTCTGTGATCCAGCGCCCTCCCTGACTATACCTGACTTCTTCTTGGCGAGTCAGCC 660

DB 601 GCGTCTGTGATCCAGCGCCCTCCCTGACTATACCTGACTTCTTCTTGGCGAGTCAGCC 660

QY 661 AGGGGACATCATGTCAGGCGGGATGA 685

DB 661 AGGGGACATCATGTCAGGCGGGATGA 685

RESULT 2

BC031665 1817 bp mRNA linear ROD 07-OCT-2003

LOCUS Mus musculus chemokine (C-X-C motif) receptor 4, mRNA (cDNA clone MGC:36266 IMAGE:3592479), complete cds.

BC031665

ACCESSION BC031665.1 GI:21618781

VERSION

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1817)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.J., Jordan, H., Moore, J., Max, S.I., Wang, J., Hsieh, P., Hopkins, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 1817)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.


```
/db_xref="SWISS-PROT:P70658"
/translation="MEPISVIVTSDNYSEVSGDYSDNSKEPCFDRDNVHNFRLP
TIYFIFLTGIVGNGLVILVMGYQKRLSMTDKYRLHLSVADLLFVITLFWAVDAMA
DWYFKFLCKAVHIYVNLVSSVLIILAFISLDRLYLAIVHATNSQPRKLLAEKAVYV
GWIIPALLITPDPFADVDSQDITSDQDRIYCDRLYPSLMMVVFQFQHIMVGLVLP
GIVILSCYCIILISKLSHSKGQRKALKTKVILILAFACWLPYVVGISDSFILLGV
IKQSCDPESIVHKWISITELAFPHCLNPLIYAFVFLGAKFKSSAQHALNSMRGSLK
ILSKRGHSSVSTESSESSPHS"
1498..1503
/gene="lestr"

polyA_signal
ORIGIN
Query Match 92.1%; Score 631.2; DB 10; Length 1575;
Best Local Similarity 99.5%; Pred. No. 9.6e-156;
Matches 633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 50 GCAGGTAGCAGTACCTCTGAGCGGTTTGGTGTCCGGTAACACCAACCGCTGTAGAGC 109
Db 20 GCAGGTAGCAGTACCTCTGAGCGGTTTGGTGTCCGGTAACACCAACCGCTGTAGAGC 79
Qy 110 GAGTGTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGA 169
Db 80 GAGTGTGCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGA 139
Qy 170 AGTGGGGTCTGGAGACTATGACTCCAAACAGGAACCTGTCTCCGGGATGAAACCGTCCA 229
Db 140 AGTGGGGTCTGGAGACTATGACTCCAAACAGGAACCTGTCTCCGGGATGAAACCGTCCA 199
Qy 230 TTTCAATAGGATCTTCTGCGCCCACTACTTCTATCATCTTCTTCTGACTGSCATAGTCGG 289
Db 200 TTTCAATAGGATCTTCTGCGCCCACTACTTCTATCATCTTCTTCTGACTGSCATAGTCGG 259
Qy 290 CAATGGATGTGTGATCTGCTGACTTGTGGAAATTTTGTGAAGGCTGTCCATATCAT 349
Db 260 CAATGGATGTGTGATCTGCTGACTTGTGGAAATTTTGTGAAGGCTGTCCATATCAT 319
Qy 350 GTACCGGTGACCTGTGAGTGGTGAACCTCTCTTGTGATCATCATCTCCCTTCTGGGC 409
Db 320 GTACCGGTGACCTGTGAGTGGTGAACCTCTCTTGTGATCATCATCTCCCTTCTGGGC 379
Qy 410 AGTTGATGCGATGCTGCTGACTTGTGGAAATTTTGTGAAGGCTGTCCATATCAT 469
Db 380 AGTTGATGCGATGCTGCTGACTTGTGGAAATTTTGTGAAGGCTGTCCATATCAT 439
Qy 470 CTACACTGTCAACTCTTACAGAGGTTCTCATCTGCGCTTATCATGAGCTGGACCGGTA 529
Db 440 CTACACTGTCAACTCTTACAGAGGTTCTCATCTGCGCTTATCATGAGCTGGACCGGTA 499
Qy 530 CCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAACTGTGGCTGAAAAGGC 589
Db 500 CCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAACTGTGGCTGAAAAGGC 559
Qy 590 AGTCTATGTGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCTTCTTGC 649
Db 560 AGTCTATGTGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCTTCTTGC 619
Qy 650 CGACCTCAGCCAGGGGACATCAGTCAGGGGATGA 685
Db 620 CGACCTCAGCCAGGGGACATCAGTCAGGGGATGA 655

RESULT 4
AB000803 1809 bp mRNA linear ROD 05-FEB-1999
LOCUS Mouse mRNA for murine CXCR-4, complete cds.
DEFINITION AB000803
ACCESSION AB000803
VERSION AB000803.1 GI:1816445
KEYWORDS murine CXCR-4; fusin.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1809)
```

```
AUTHORS Suzuki, G., Nakata, Y., Uzawa, A., Shirasawa, T., Saito, T. and Mita, K.
TITLE Molecular cloning of murine CXCR-4, a murine homologue of HIV entry
co-receptor, that is expressed on leukocyte and other organs
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1809)
AUTHORS Suzuki, G.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-1997) Gen Suzuki, National Institute of
Radiological Sciences, Division of Radiation Health; 4-9-1,
Anagawa, Inage-ku, Chiba, Chiba 263, Japan
(E-mail: gsuzuki@nmx1.nirs.go.jp, Tel: 043-251-2111,
Fax: 043-284-1736)
FEATURES
Location/Qualifiers
1..1809
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL-6"
/db_xref="taxon:10090"
/tissue_type="thymus"
/dev_stage="E17"
72..1151
/note="fusin"
/codon_start=1
/product="murine CXCR-4"
/protein_id="BAA19187.1"
/db_xref="GI:1816446"
/translation="MEPISVIVTSDNYSEVSGDYSDNSKEPCFDRDNVHNFRLP
TIYFIFLTGIVGNGLVILVMGYQKRLSMTDKYRLHLSVADLLFVITLFWAVDAMA
DWYFKFLCKAVHIYVNLVSSVLIILAFISLDRLYLAIVHATNSQPRKLLAEKAVYV
GWIIPALLITPDPFADVDSQDITSDQDRIYCDRLYPSLMMVVFQFQHIMVGLVLP
GIVILSCYCIILISKLSHSKGQRKALKTKVILILAFACWLPYVVGISDSFILLGV
IKQSCDPESIVHKWISITELAFPHCLNPLIYAFVFLGAKFKSSAQHALNSMRGSLK
ILSKRGHSSVSTESSESSPHS"
ORIGIN
Query Match 90.3%; Score 618.6; DB 10; Length 1809;
Best Local Similarity 99.2%; Pred. No. 2.1e-152;
Matches 632; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Qy 50 GCAGGTAGCAGTG-ACCTCTGAGCGGTTTGGTGTCCGGTAACACCAACCGCTGTAGAG 108
Db 1 GCAGGTAGCAGTGAACCTCTGAGCGGTTTGGTGTCCGGTAACACCAACCGCTGTAGAG 60
Qy 109 CGAGTGTTCCTCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAG 168
Db 61 CGAGTGTTCCTCCATGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAG 120
Qy 169 AAGTGGGGTCTGGAGACTATGACTCCAAACAGGAACCTGTCTCCGGGATGAAAACGTC 228
Db 121 AAGTGGGGTCTGGAGACTATGACTCCAAACAGGAACCTGTCTCCGGGATGAAAACGTC 180
Qy 229 ATTTCAATAGGATCTTCTGCGCCACCATCTTCTCATCATCTTCTTGAAGCTAGTCG 288
Db 181 ATTTCAATAGGATCTTCTGCGCCACCATCTTCTCATCATCTTCTTGAAGCTAGTCG 240
Qy 289 GCAATGGATTGGTGTGATCTCTGGTTCATGGGTACCAAGAAAGCTAAGGAGCATGACGACA 348
Db 241 GCAATGGATTGGTGTGATCTCTGGTTCATGGGTACCAAGAAAGCTAAGGAGCATGACGACA 300
Qy 349 AGTACCGGTGACCTGTGAGTGGCTGACTCTTCTTGTGATCATCATCTCCCTTCTGG 408
Db 301 AGTACCGGTGACCTGTGAGTGGCTGACTCTTCTTGTGATCATCATCTCCCTTCTGG 360
Qy 409 CAGTTCATGCCATGGCTGACTGTGACTTTTGGGAAATTTTGTGAAGGCTGTCCATATCA 468
Db 361 CAGTTCATGCCATGGCTGACTGTGACTTTTGGGAAATTTTGTGAAGGCTGTCCATATCA 420
Qy 469 TCTACACTGTCAACTCTTACAGAGGTTCTCATCTGCGCTTCTATGAGCTGGACCGGT 528
Db 421 TCTACACTGTCAACTCTTACAGAGGTTCTCATCTGCGCTTCTATGAGCTGGACCGGT 480
Qy 529 ACCTCGCCATTGTCCACGCCACCAACAGTCAAGGCCAAGGAACTGTGGCTGAAAAGC 588
```

Suzuki, G., Nakata, Y., Uzawa, A., Shirasawa, T., Saito, T. and Mita, K.
Molecular cloning of murine CXCR-4, a murine homologue of HIV entry
co-receptor, that is expressed on leukocyte and other organs
Unpublished
2 (bases 1 to 1809)
Suzuki, G.
Direct Submission
Submitted (29-JAN-1997) Gen Suzuki, National Institute of
Radiological Sciences, Division of Radiation Health; 4-9-1,
Anagawa, Inage-ku, Chiba, Chiba 263, Japan
(E-mail: gsuzuki@nmx1.nirs.go.jp, Tel: 043-251-2111,
Fax: 043-284-1736)

Db 481 ACCTCGCTATTGTCACGCCACCAACAGTCAGAGGCCAAGAACTGCTGGCTGAAAGG 540
Qy 589 CAGTCTATGCGGCTGTGATCCGAGCCCTCTCTGACTATACCTGACTTCACTTTTG 648
Db 541 CAGTCTATGCGGCTGTGATCCGAGCCCTCTCTGACTATACCTGACTTCACTTTTG 600
Qy 649 CCGAGCTCAGCCAGGGGAGCATCAGTCAGGGGGATGA 685
Db 601 CCGAGCTCAGCCAGGGGAGCATCAGTCAGGGGGATGA 637

RESULT 5
MMU59760 1223 bp mRNA linear ROD 30-SEP-2002
LOCUS Mus musculus fusin homolog mRNA, complete cds.
ACCESSION U59760
VERSION U59760.1 GI:1527134
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1223)
AUTHORS Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1
co-factor
JOURNAL J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE 97113334
PubMed 8955194
REFERENCE 2 (bases 1 to 1223)
AUTHORS Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA
FEATURES
source 1..1223
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv"
/db_xref="taxon:10090"
30..1109
/note="seven transmembrane spanning receptor;
chemokine-receptor like"
/codon_start=1
/product="fusin homolog"
/protein_id="AA807725.1"
/db_xref="GI:1527135"
/translations="MEPISVSVITSDNYSEVSGDYDSNKEPCFRDENVHFNRIPLP
TIYFIPLTGIVGNGLVIMVGYQKLRMTDKYRLHLSVADLLFVITLFFWADAMA
DWYFGLCAVHTIYTWLYSSVLLARISLDRYLAVHTANSQSPKLLAEKAVYV
GVNIPALLIIPFIADVSQDISQGDRIYCDRLYPSLKNVWFQFHWGLILP
GVILSCYCIILSKLSHGKQRKALKITVILILAFFACWLEFYVIGISIDSLILGV
IKQCDPESIVHWKISITELAFPHCLNPLIYAFILGAKFKPSAQHALNSMRGSLK
ILSKGRGHHSSVSTESSFHS"

ORIGIN
Query Match 86.9%; Score 595; DB 10; Length 1223;
Best Local Similarity 100.0%; Pred. No. 3.5e-146; Gaps 0;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 91 ACCACCAAGGCTGTAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATACACTT 150
Db 1 ACCACCAAGGCTGTAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATACACTT 60
Qy 151 CTGATACTACTCTGAGAAGTGGGCTCGAGACTATGACTCAACAGGACCCCTGCT 210
Db 61 CTGATACTACTCTGAGAAGTGGGCTCGAGACTATGACTCAACAGGACCCCTGCT 120
Qy 211 TCCGGGATGAAACGTCATTCATAGGATCTTCCTGCCACCATCTACTTTCATCATCT 270
Db 121 TCCGGGATGAAACGTCATTCATAGGATCTTCCTGCCACCATCTACTTTCATCATCT 180
Qy 271 TCTTGACTGCGCATAGTCGGCAATGGATTGGTGATCTCTGGTCATGGGTACCAGAGAAGC 330

Db 181 TCTTGACTGGCATAGTCGGCAATGGAATGGTGATCTCTGGTTCATGGTTACCAAGAAGC 240
Qy 331 TAAGAGCATGACGGAACAAGTACCGGCTGCACCTGTGAGTGGTGCAGTCTCTTTCTCA 390
Db 241 TAAGAGCATGACGGAACAAGTACCGGCTGCACCTGTGAGTGGTGCAGTCTCTTTCTCA 300
Qy 391 TCACACTCCCCCTTCTGGGCAAGTTCATGCCATGGCTGACTTGGTACTTTGGGAAATTTTGT 450
Db 301 TCACACTCCCCCTTCTGGGCAAGTTCATGCCATGGCTGACTTGGTACTTTGGGAAATTTTGT 360
Qy 451 GTAAGGCTGTCAT 510
Db 361 GTAAGGCTGTCAT 420
Qy 511 TCATCAGCTGACGCGGCTACCTCCGCAATGTCACGACCCACCAACAGTCAAGGCAAGGA 570
Db 421 TCATCAGCTGACGCGGCTACCTCCGCAATGTCACGACCCACCAACAGTCAAGGCAAGGA 480
Qy 571 AACTGCTGGCTGAAAAGGCAAGTCTATATGCGGCTGTGATCCAGCCCTCTCTCTGACTA 630
Db 481 AACTGCTGGCTGAAAAGGCAAGTCTATGCGGCTGTGATCCAGCCCTCTCTCTGACTA 540
Qy 631 TACCTGACTTTCATCTTTGCGGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 685
Db 541 TACCTGACTTTCATCTTTGCGGACGTCAGCCAGGGGACATCAGTCAGGGGATGA 595

RESULT 6
MMLCR12 1180 bp mRNA linear ROD 18-NOV-1997
LOCUS Mus musculus lcr-1 gene.
DEFINITION Z80112
ACCESSION Z80112.1 GI:2632100
VERSION CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1180)
AUTHORS Schubel, A., Burgstahler, R. and Lipp, M.
TITLE The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
High expression in thymus and lymphoid tissues
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1180)
AUTHORS Lipp, M.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
Molecular Medicine, MDC, Robert-Rössle-Strasse 10, BERLIN-BUCH,
D-13122, GERMANY
COMMENT On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES
source location/Qualifiers
1..1180 /organism="Mus musculus"
/mol_type="mRNA"
/strain="129SV"
/db_xref="taxon:10090"
/tissue_type="thymus"
/dev_stage="adult"
1..1180 /gene="lcr-1"
1..1074 /gene="lcr-1"
/function="receptor for chemokine SDF-1"
/standard_name="CXC chemokine receptor 4"
/codon_start=1
/product="CXCR-4"
/protein_id="CAB02202.1"
/db_xref="GI:1542889"
/db_xref="GOA:P70658"
/db_xref="SWISS-PROT:P70658"
/translation="MEPISVSVITSDNYSEVSGDYDSNKEPCFRDENVHFNRIPLPTI
YFIPLTGIVGNGLVIMVGYQKLRMTDKYRLHLSVADLLFVITLFFWADAMADW

MMU65580
LOCUS 3366 bp DNA linear ROD 13-DEC-1996
DEFINITION Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION U65580
VERSION U65580.1 GI:1731650
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE 1 (bases 1 to 3366)
AUTHORS Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1 co-factor
JOURNAL J. Immunol. 157 (12), 5455-5460 (1996)
MEDLINE 97113334
PUBMED 8955194
REFERENCE 2 (bases 1 to 3366)
AUTHORS Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA
FEATURES
source
1. .3366
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1. .3366
/gene="CXCR-4"
join(1..21,2308..3366)
/gene="CXCR-4"
/note="SDP-1 chemokine receptor"
/codon_start=1
/product="fusin"
/protein_id="AAC52953.1"
/db_xref="GI:3645915"
translations="MPSIVSYTSDNVEEVSQDYSNKGPCFRDNVFNRIPLP
TIYFIFLGVNGVILVMGYOKLRSMIDKRLHLSVADLLFVITLPFAVDAMA
DWYFGLCKAVHIYVNLVSVLJLAPISLDRVLAVHATNSORPKLAEKAVVY
GVWIPALLTIPFIADVSQDSDYICDRILPDLNWMVQFQHWGLILP
GIVLSCYIIISKLSHSGHQRKALKTTVILILAFACWLPYVIGSIDSFILIGV
IKQCDPFIHVKWISITLALAPFHCHLNPILYFLGAKFKSAQHNLNMSRSGSLK
ILSKRGKGGHSSVSTSESSPHSS"
1. .21
/gene="CXCR-4"
/number=1
1. .15
/gene="CXCR-4"
/note="alternate exon 1"
16. .>2307
/gene="CXCR-4"
/note="alternate intron 1"
22. .2307
/gene="CXCR-4"
/number=1
2308. .3366
/gene="CXCR-4"
/number=2

Query Match 79.7%; Score 545.8; DB 10; Length 3366;
Best Local Similarity 97.9%; Pred. No. 3.7e-133;
Matches 553; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 121 TCGAACCGATCAGTGTGAGTATACACTTCTGATACCTCTGAGAAAGTGGGCTCTG 180
DB 2288 TGTCCTCTGTTCTTTCAGATATACACTTCTGATACCTCTGAGAAAGTGGGCTCTG 2347
QY 181 GAGACTATGACTCCAAACAGAAACCTGCTTCGCGGATGAAACAGTCCATTTCAATAGGA 240
DB 2348 GAGACTATGACTCCAAACAGAAACCTGCTTCGCGGATGAAACAGTCCATTTCAATAGGA 2407

QY 241 TCTTCTGCGCCACCATCTACTTCAATCATCTTCTTGTGACTGGCATAGTCGGCAATGATGG 300
DB 2408 TCTTCTGCGCCACCATCTACTTCAATCATCTTCTTGTGACTGGCATAGTCGGCAATGATGG 2467
QY 301 TGATCCTGGTTCATGGTTTACCAGAGAGAGCTAAGAGAGATGACGACGACAAAGTACCGGCTGC 360
DB 2468 TGATCCTGGTTCATGGTTTACCAGAGAGAGCTAAGAGAGATGACGACGACAAAGTACCGGCTGC 2527
QY 361 ACCTGTGCTAGTGGTGCACCTCTCTTGTGATCATCACATCCCTTCTGGGCAAGTTCATGCCA 420
DB 2528 ACCTGTGCTAGTGGTGCACCTCTCTTGTGATCATCACATCCCTTCTGGGCAAGTTCATGCCA 2587
QY 421 TGGCTGACTGCTACTTTGGGAAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 480
DB 2588 TGGCTGACTGCTACTTTGGGAAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCA 2647
QY 481 ACCTCTACAGCAGCGTCTCTCATCTCGCTTTCATCAGCTGACCGGTACCTGCCATTG 540
DB 2648 ACCTCTACAGCAGCGTCTCTCATCTCGCTTTCATCAGCTGACCGGTACCTGCCATTG 2707
QY 541 TCCAGGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGGAAGGCAAGTCTATGTGG 600
DB 2708 TCCAGGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGGAAGGCAAGTCTATGTGG 2767
QY 601 CGGTCTGGATCCAGCCCTCTCTCTGACTACTACTGACTTCTTCTTCCGACGCTCAGCC 660
DB 2768 CGGTCTGGATCCAGCCCTCTCTCTGACTACTACTGACTTCTTCTTCCGACGCTCAGCC 2827
QY 661 AGGGGAGATCAGTCAGGGGATGA 685
DB 2828 AGGGGAGATCAGTCAGGGGATGA 2852

RESULT 9
MMLESTRGN 3770 bp DNA linear ROD 02-APR-1997
LOCUS M.musculus gene encoding leukocyte-derived seven transmembrane domain receptor, strain B6.
DEFINITION
ACCESSION X99581
VERSION X99581.1 GI:1924959
KEYWORDS lestr gene; leukocyte-derived seven transmembrane domain receptor.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Moepps, B., Frodl, R., Kessler, H. and Gierschik, P.
TITLE cDNA cloning and genomic organization of a leukocyte-derived seven transmembrane domain receptor (LESTR) from mouse: a murine homologue of the human HIV-1 entry cofactor fusin
JOURNAL Unpublished
REFERENCE 2
AUTHORS Moepps, B.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm, Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
REMARK Revised by [3]
REFERENCE 3 (bases 1 to 3770)
AUTHORS Moepps, B.
TITLE Direct Submission
JOURNAL Submitted (02-APR-1997) B. Moepps, Universitaet Ulm, Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG
COMMENT On Apr 3, 1997 this sequence version replaced gi:1666646.
FEATURES
source
1. .3770
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="B6"
/db_xref="taxon:10090"
/haplotype="diploid"
/tissue_type="thymus"
/dev_stage="6-8 weeks old"
1. .110
exon

* 1	contig of 5345 bp in length
* 5345	contig of 5345 bp in length
* 5445	gap of unknown length
* 5446	contig of 11750 bp in length
* 17195	contig of 11750 bp in length
* 17196	gap of unknown length
* 17296	contig of 15955 bp in length
* 33251	gap of unknown length
* 33251	gap of unknown length
* 33251	contig of 12296 bp in length
* 45646	gap of unknown length
* 45647	gap of unknown length
* 45747	contig of 16246 bp in length
* 61992	contig of 16246 bp in length
* 61993	gap of unknown length
* 63093	contig of 1331 bp in length
* 63243	contig of 1331 bp in length
* 63424	gap of unknown length
* 63523	gap of unknown length
* 64566	contig of 1143 bp in length
* 64676	gap of unknown length
* 64677	gap of unknown length
* 65994	contig of 1228 bp in length
* 65995	gap of unknown length
* 66934	gap of unknown length
* 66935	contig of 1336 bp in length
* 67431	gap of unknown length
* 67530	gap of unknown length
* 67531	contig of 1452 bp in length
* 68982	contig of 1452 bp in length
* 69082	gap of unknown length
* 69083	contig of 1809 bp in length
* 70991	gap of unknown length
* 70992	gap of unknown length
* 70992	contig of 1181 bp in length
* 72172	contig of 1181 bp in length
* 72173	gap of unknown length
* 72272	gap of unknown length
* 73614	contig of 1342 bp in length
* 73714	gap of unknown length
* 75185	contig of 1471 bp in length
* 75186	gap of unknown length
* 75286	contig of 1569 bp in length
* 76855	gap of unknown length
* 76855	contig of 1459 bp in length
* 78413	contig of 1459 bp in length
* 78513	gap of unknown length
* 78514	contig of 1451 bp in length
* 79364	contig of 1451 bp in length


```
* 7995 80064: gap of unknown length
* 80065 81745: contig of 1681 bp in length
* 81746 81845: gap of unknown length
* 81846 83417: contig of 1572 bp in length
* 83418 83517: gap of unknown length
* 83518 84910: contig of 1393 bp in length
* 84911 85010: gap of unknown length
* 85011 86788: contig of 1778 bp in length
* 86789 86888: gap of unknown length
* 86889 88992: contig of 2004 bp in length
* 88993 90670: gap of unknown length
* 90671 90770: contig of 1678 bp in length
* 90771 93207: contig of 2437 bp in length
* 93208 93308: gap of unknown length
* 93309 96559: contig of 3351 bp in length
* 96560 96759: gap of unknown length
* 96760 100577: contig of 3818 bp in length
* 100578 104988: gap of unknown length
* 104989 105089: contig of 4312 bp in length
* 105090 108993: gap of unknown length
* 108994 112633: contig of 3639 bp in length
* 112634 116328: gap of unknown length
* 116329 116429: contig of 3596 bp in length
* 116430 120980: gap of unknown length
* 120981 124080: contig of 4552 bp in length
* 124081 126821: contig of 5741 bp in length
* 126822 131331: contig of 4410 bp in length
* 131332 131432: gap of unknown length
* 131433 138706: contig of 7274 bp in length
* 138707 145169: gap of unknown length
* 145170 145269: contig of 6364 bp in length
* 145270 150698: gap of unknown length
* 150699 150797: gap of unknown length
* 150798 159989: contig of 9192 bp in length
* 159990 160089: gap of unknown length
* 160090 167244: contig of 7156 bp in length
* 167245 167345: gap of unknown length
* 167346 175370: contig of 8026 bp in length.
FEATURES
    Location/Qualifiers
        1..175370
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /chromosome="UNK"
            /clone="RP43-21D9"
        1..5345
            /note="assembly_name:Contig100"
        5446..17195
            /note="assembly_name:Contig101"
        17296..33250
            /note="assembly_name:Contig102"
        33351..45646
            /note="assembly_name:Contig103"
        45747..61992
            /note="assembly_name:Contig104"
        62093..63423
            /note="assembly_name:Contig16"
        63524..64666
            /note="assembly_name:Contig3"
        64767..65994
            /note="assembly_name:Contig45"
        66095..67430
            /note="assembly_name:Contig48"
        67531..68982
            /note="assembly_name:Contig5"
        69083..70891
            /note="assembly_name:Contig56"
        70992..72172
            /note="assembly_name:Contig96"
        72273..73614
            /note="assembly_name:Contig97"
        73715..75185
            /note="assembly_name:Contig98"
        75286..76854
            /note="assembly_name:Contig99"
        76955..78413
            /note="assembly_name:Contig73"
        78514..79964
            /note="assembly_name:Contig75"
        80065..81745
            /note="assembly_name:Contig77"
        81846..83417
            /note="assembly_name:Contig79"
        83518..84910
            /note="assembly_name:Contig80"
        85011..86788
            /note="assembly_name:Contig81"
        86889..88992
            /note="assembly_name:Contig82"
        88993..90670
            /note="assembly_name:Contig83"
        90771..93207
            /note="assembly_name:Contig84"
        93308..96559
            /note="assembly_name:Contig85"
        96760..100577
            /note="assembly_name:Contig86"
        100678..104988
            /note="assembly_name:Contig87"
        105089..108993
            /note="assembly_name:Contig88"
        108994..112632
            /note="assembly_name:Contig89"
        112733..116328
            /note="assembly_name:Contig90"
        116429..120980
            /note="assembly_name:Contig91"
        121081..126821
            /note="assembly_name:Contig92"
        126922..131331
            /note="assembly_name:Contig93"
        131432..138705
            /note="assembly_name:Contig94"
        138806..145169
            /note="assembly_name:Contig95"
        145270..150696
            /note="assembly_name:Contig96"
        150797..159988
            /note="assembly_name:Contig97"
        160089..167244
            /note="assembly_name:Contig98"
        167345..175370
            /note="assembly_name:Contig99"
        ORIGIN
            Query Match 79.0%; Score 541; DB 2; Length 175370;
            Best Local Similarity 97.3%; Pred. No. 8.5e-132;
            Matches 550; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
            QY 121 TGGACCGATCAGTGTGAGTATATACCTCTGATACTCTCTGAAGAGTGGGCTG 180
            DB 167962 TGTCCCTGTTTCCTTTCAGATATACCTCTGATACTCTCTGAAGAGTGGGCTG 168021
            QY 181 GAGACTATGACTCCCAAGGAACCCCTGCTCCGGATGAAACGTCATTTCATAGGA 240
            DB 168022 GAGACTATGACTCCCAAGGAACCCCTGCTCCGGATGAAACGTCATTTCATAGGA 168081
            QY 241 TCTTCTGCCACATCTTCTTCATCTCTTCTGCGCATAGTCGCGCATGATGG 300
            DB 168082 TCTTCTGCCACATCTTCTTCATCTCTTCTGCGCATAGTCGCGCATGATGG 168141
```



```

/protein_id="CAB02201.1"
/db_xref="GI:1542887"
/db_xref="GOA:P70658"
/db_xref="SWISS-PROT:P70658"
/translation="MPESTIVTSNYSSEVGSDYSNKEPCFRDENVHNFELFPTI
YPIELFLGIVGNGLVILPMGQKLRSMNTDKYRLHLSVADLLFVITLPFWADVADADW
YFGFLCKARAHIIIVTNDYSSVLLIAFLSLDRYLAIIVHATNSQRPR"
99..>528
/gene="lcr-1"
/number=2
exon

```

230	CAATGGATTGGTGAATCTCTGGTCAATGGGTTACAGAAAGCACTAAGGAGCATACGGACAA	349
248	CAATGGATTGGTGAATCTCTGGTCAATGGGTTACAGAAAGCACTAAGGAGCATACGGACAA	307
350	GTACGGGCTGCACCTGTCAAGTGGCTGACCTCTCTTTGTTCATCACACTCCCTTCTGGGC	409
368	GTACGGGCTGCACCTGTCAAGTGGCTGACCTCTCTTTGTTCATCACACTCCCTTCTGGGC	367
410	AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCAT	469
428	AGTTGATGCCATGGCTGACTGGTACTTTGGGAAATTTTGTGTAGGCTGTCCATATCAT	427
470	CTACACTGTCAACTCTACAGAGCGTTCTCATCTCGCCCTTCATCAGCCTGGACGGTA	529
488	CTACACTGTCAACTCTACAGAGCGTTCTCATCTCGCCCTTCATCAGCCTGGACGGTA	487
530	CCTGCCATTCTCACAGCCACCAACAGTCAAGGCCAAGGA	570
548	CCTGCCATTCTCACAGCCACCAACAGTCAAGGCCAAGGA	528

AC122097 190149 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-108A24, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC122097
AC122097.4 GI:25008263
HTG; HTGS PHASE1; HTGS RATT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1. (bases 1 to 190149)
Murny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allien,H., Alabrooks,S., Amin,A., Anguiano,D.,
Avalyabechi,V., Ayogei,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Behnam,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.

QY	256	TCTACTCATGATCATTCTTGTAC	TGGCATAGTCGGCAATGGATTTGGTGATCCTCGGTGCATGG	315	
Db	122	TCTATTTTTATCATCTCTTTGACT	TGCGCATAGTCGGCAATGGGTTGGTAATCCTTGGTGCATGG	181	
QY	316	GTTTACCAGAAGAAGCTTAAGGAG	CATCACGGACAAGTACCCGGTGCACCTGTCTCAGTGGCTG	375	
Db	182	GTTTACCAGAAGAAGCTGAGGAG	CATCACGAACAAGTACCGGCTGCACCTGTCTCGTGGCTG	241	
QY	376	ACTCTCTCTTTGTATCATCACTCT	CCCCCTTCTGGGCAGTTGATGCCATGGCTGACTGGTACT	435	
Db	242	ACCTCTCTTTGTATCATCACTCCC	CTTCTGGGCAGTTGATGCCATGGCTGACTGGTACT	301	
QY	436	TTTGGGAAATTTTGTGTGAAGGCT	GTCCATCATCTATCATCTGTCAAACCTCTACAGCAGCG	495	
Db	302	TTTGGGAAATTTTGTGAAGGCTGT	GCATATCATCTACACCGTCAAACCTTTACAGCAGTG	361	
QY	496	TTCTCATCTCGGCTTTCATCAGCT	TGGACGGTACCTCGCCATTTGCCACGCCACCAACA	555	
Db	362	TTCTCATCTCGGCTTTCATCAGCT	TGGACGGTACCTCGCCATTTGCCACGCCACCAACA	421	
QY	556	GTCMAAGGCCAAGGAACTGTGGCT	GTAAAAGGCAGCTATATGTGGGGCTCTGGATCCCGAG	615	
Db	422	GCCAGAGGCCGAGGAAGCTGTGGC	TGTAAAAGGCCTCTATGTGGGTCTCTGGATCCCGC	481	
QY	616	CCCTCTCTCTGACTATACCTGACT	TTCATCTTTGCCGAGCTCAGCCAGGGGGACATCAG	673	
Db	482	CCCTCTCTCTGACTATACCTGACT	TTCATCTTTGCCGAGCTCAGCCAGGGGGACGGCAG	539	
RESULT 15	RNU90610	1050 bp	mRNA	linear	ROD 26-MAR-1997
LOCUS	Rattus norvegicus CXc chemokine receptor (CXCR4)				mRNA, complete cds.
DEFINITION	U90610.1	GI:1906612			
ACCESSION	Rattus norvegicus (Norway rat)				
VERSION	Rattus norvegicus				
KEYWORDS	Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
SOURCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
ORGANISM	Rattus				
REFERENCE	1 (bases 1 to 1050)				
AUTHORS	Harrison,J.K. and Salafranca,M.N.				
TITLE	Molecular cloning of rat CXCR4				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1050)				
AUTHORS	Harrison,J.K. and Salafranca,M.N.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-FEB-1997) Pharmacology and Therapeutics, University of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA				
FEATURES	Location/Qualifiers				
source	1..1050				
	/organism="Rattus norvegicus"				
	/mol_type="mRNA"				
	/strain="Wistar"				
	/db_xref="taxon:10116"				
	/tissue_type="spleen"				
	1..1050				
	/gene="CXCR4"				
gene	1..1050				
	/gene="CXCR4"				
CDS	/notes="G-protein coupled receptor"				
	/codon_start=1				
	/product="CXc chemokine receptor"				
	/protein_id="AAB50408.1"				
	/db_xref="GI:1906613"				
	/translation="MLVGYSDNSVEVGSDYDNKEPCFRDENENFPLPTIYFI IFLFTGVGLVTVLVGYOKFLASMTDKVLHLSVDLLFLVTLPFAWDADMYGF KFLCKAVHIITYNLVSUVILAFISLDRLVAIVHATNSQARKLAELAEKAYVCWVIE ALLTLTDPIFDVDSODGRIGIDRLPYSLSNVVQFOHQINVLILPGVILLSCYCI ITSKLSHGSHQKKALKTTVILILFAFWLPPYYGVGISIDSFILLEVILKQCCEFS				

VHKWISLTELAFPHCLNPLIYAFGLGAKFKSSAQHALNSMRSGSLKILSKGKRGGH
SSVSTSESSSFHSS"

ORIGIN

Query Match	68.5%;	Score	469.2;	DB	10;	Length	1050;
Best Local Similarity	92.0%;	Pred. No.	6.3e-113;				
Matches	495;	Conservative	0;	Mismatches	43;	Indels	0;
Gaps	0;						

Qy	136	TGAGTATATACACTCTCTGATTAACCTACTCTGAAGAGTGGGTCTGGAGACTATGACTCCA	195
Db	2	TGGAAATATACACTCTGGATTAACCTACTCTGGAAGAGTAGGTCTGGAGACTATGACTCCA	61
Qy	196	ACAAGGAACCTCTTCGGGATCAAAACGTCCTCAATTAAGGATCTTCTGCCCCACCA	255
Db	62	ACAAGGAACCTCTTCGGGATCAAAACGTCCTCAATTAAGGATCTTCTGCCCCACCA	121
Qy	256	TCTACTTCATCATCTTCTGCTGCTAGTCGGCAATGGATTGGTGTGATCTTGTGTATGG	315
Db	122	TCTATTTTATCATCTTCTGCTGCTAGTCGGCAATGGATTGGTGTGATCTTGTGTATGG	181
Qy	316	GTTACCAAGAAGCTAAGGAGCATGACGGAACAAGTACCGGCTGCACCTGTCTAGTGGCTG	375
Db	182	GTTACCAAGAAGCTAAGGAGCATGACGGAACAAGTACCGGCTGCACCTGTCTAGTGGCTG	241
Qy	376	ACCTCCTCTTTGTCTATCACACTCCCTCTCGGGAGTTGATGCCATGGCTGACTGGTACT	435
Db	242	ACCTCCTCTTTGTCTATCACACTCCCTCTCGGGAGTTGATGCCATGGCTGACTGGTACT	301
Qy	436	TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCG	495
Db	302	TTGGGAAATTTTGTGTAAGGCTGTGCTATCATCTACCCGTCAACCTTTACAGCAGTG	361
Qy	496	TTCTCATCTGGCCTTCATACAGCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACA	555
Db	362	TTCTCATCTGGCCTTCATACAGCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACA	421
Qy	556	GTCRAAGGCCAAGGAACCTGTGGCTGAAAGGAGCTCTATGTGGGGTCTGGATCCAG	615
Db	422	GCCAGCGCGAGGAAGCTGTGGCTGAAAGGCGCTATGTGGGTGTCTGGATCCCG	481
Qy	616	CCCTCCTCCTGACTATACCTGACTTTCATCTTTGCCGAGTCAGCCAGGGGACATCAG	673
Db	482	CCCTCCTCCTGACTATCCTGACTCATCTTTCGCCGATGTCCAGCCAGGGGACGGCAG	539

Search completed: May 17, 2004, 12:10:47
Job time : 2772.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 03:46:58 ; Search time 293.749 Seconds

(Without alignments)
9906.453 Million cell updates/sec

Title: US-09-367-052-5

Perfect score: 685

Sequence: 1 ccatacctaatacgaactcact.....gacatcagtcagggggatga 685

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	685	100.0	1877	2	Aav46370 Nucleic a
2	566	82.6	1758	2	Aaz27611 Mouse CX
3	469.2	68.5	1050	9	AdB58421 Toxicity-
4	469.2	68.5	1050	9	AdB53005 Primary r
5	446.8	65.2	1679	4	Aad12801 Human neu
6	446.8	65.2	1679	4	Aai65467 Nucleotid
7	446.8	65.2	1679	6	AbS53992 DNA encod
8	446.8	65.2	1679	6	AbS53750 DNA encod
9	446.8	65.2	1679	6	Abv78039 Hypoxia-r
10	446.8	65.2	1679	7	Abx08779 Argiogene
11	446.8	65.2	1679	7	Abx74454 Human CDN
12	446.8	65.2	1679	7	Abz68886 Nucleotid
13	446.8	65.2	1679	7	Abz42642 Human CX
14	446.8	65.2	1679	9	AdC98645 Human CX
15	446.8	65.2	1711	6	Abz55348 Human gen
16	446.8	65.2	1711	9	AdB47320 Human CX
17	445.8	65.1	1664	3	AAAG1656 Human CX
18	445.8	65.1	1664	3	AAZ40014 CXCR4 cod
19	439.2	64.1	1737	2	AAQ29506 New plate
20	439.2	64.1	1737	2	AAQ80521 Human mon
21	439.2	64.1	1737	2	AAQ90007 Chemokine
22	437.6	63.9	1670	6	ABK3803 Human cDN
23	437.6	63.9	1670	6	ABN95645 Gene #214

24	437.6	63.9	1670	7	ACC46765	ACC46765 Human COP
25	432.4	63.1	1102	7	ACA64750	ACA64750 Chemokine
26	430.8	62.9	1059	3	AAZ38553	AAZ38553 Human CX
27	430.8	62.9	1059	5	AB197982	AB197982 Non-endog
28	430.8	62.9	1059	6	ABQ74938	ABQ74938 Human CX
29	430.8	62.9	1059	6	AAH99952	AAH99952 CXCR4 enc
30	430.8	62.9	1119	6	AAK98241	AAK98241 Human cys
31	430.8	62.9	1225	6	ABZ35630	ABZ35630 Human gen
32	430.8	62.9	1225	7	ACA56637	ACA56637 Human sig
33	430.8	62.9	1588	2	AAZ27610	AAZ27610 Human CX
34	429.2	62.7	1059	7	ACC72707	ACC72707 Human can
35	428	62.5	1317	2	AAQ66179	AAQ66179 Seven tra
36	428	62.5	1317	2	AAV18357	AAV18357 Human rM3
37	428	62.5	1317	3	AAA91726	AAA91726 Human 7TM
38	428	62.5	1317	6	ABK54255	ABK54255 Human 7 t
39	428	62.5	1944	2	AAK15882	AAK15882 CDNA enco
40	428	62.5	5161	3	AAA34774	AAA34774 Human ade
41	428	62.5	5161	3	AAF20896	AAF20896 Human CCR
42	428	62.5	5161	6	AAH99951	AAH99951 CXCR4 gen
43	428	62.5	5161	7	ABZ96590	ABZ96590 Human CCR
44	428	62.5	5161	7	ACA64841	ACA64841 Human CX
45	427.6	62.4	1582	6	ABS64793	ABS64793 CXCR4 rec

ALIGNMENTS

RESULT 1

AAV46370

ID AAV46370 standard; CDNA to mRNA; 1877 BP.

AC AAV46370;

DT 20-NOV-1998 (first entry)

DB Nucleic acid encoding a murine CX chemokine receptor.

KW Mouse; CX chemokine receptor; pre-B cell line DW34;

KW CX chemokine pre-B cell stimulatory factor PBSP/SDF-1; HIV infection;

KW screening; inhibitor; AIDS; ds.

OS Mus sp.

FN Key Location/Qualifiers

FT CDS 120..1199

FT FT /*tag= a

PN WO9835035-A1.

PD 13-AUG-1998.

PF 07-FEB-1997; 97WO-JP000299.

PR 07-FEB-1997; 97WO-JP000299.

PA (SHIO) SHIONOGI & CO LTD.

PI Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;

PI Nakajima T, Yoshie O;

PI WPI; 1998-447232/38.

PI P-PSDB; AAW64778.

PT Mouse CX chemokine receptor binding to PBSP/SDF-1 pre-B cell stimulatory factor - is useful for screening of potential HIV infection and AIDS inhibitors.

PT Claim 3; Page 39-42; 76pp; Japanese.

CC The present sequence encodes a murine CX chemokine receptor which binds to the mouse CX chemokine pre-B cell stimulatory factor PBSP/SDF-1. The nucleic acid is isolated from mouse pre-B cell line DW34. The receptor and cells expressing it can be used in the study and mapping of the

CC	mechanism of HIV infection and in screening of potential inhibitors of
CC	HIV infection and the development of AIDS
XX	
QQ	Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;
Query Match	100.0%; Score 685; DB 2; Length 1877;
Best Local Similarity	100.0%; Pred. No. 5.8e-191;
Matches 685; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CCATCCTAATACGACTCAGCTATAGGGCTCGAGCGCGCCGCGGCGAGGTGCAGGTAGCAG 60
DB	1 CCATCCTAATACGACTCAGCTATAGGGCTCGAGCGCGCCGCGGCGAGGTGCAGGTAGCAG 60
QY	61 TGAACCTCTGAGCGGTTTGGTGTCCGGTAAACACACGCGCTGTAGAGCGAGTGTGCGCA 120
DB	61 TGAACCTCTGAGCGGTTTGGTGTCCGGTAAACACACGCGCTGTAGAGCGAGTGTGCGCA 120
QY	121 TGGAAACCGATCAGTGTGAGTATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTG 180
DB	121 TGGAAACCGATCAGTGTGAGTATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTG 180
QY	181 GAGACTATGACTCCAAACAGAAACCTGTCTCCGGATGAAAACGTCCATTTCAATAGGA 240
DB	181 GAGACTATGACTCCAAACAGAAACCTGTCTCCGGATGAAAACGTCCATTTCAATAGGA 240
QY	241 TCTTCTGCGCCACCATCTACTTTCATCATCTTCTTGACTGTCATAGTCGGCAATGGATTGG 300
DB	241 TCTTCTGCGCCACCATCTACTTTCATCATCTTCTTGACTGTCATAGTCGGCAATGGATTGG 300
QY	301 TGATCTGTGTCATGGGTTACCAAGAAGCTAAGAGCATGACGCAAGTACCGGCTGC 360
DB	301 TGATCTGTGTCATGGGTTACCAAGAAGCTAAGAGCATGACGCAAGTACCGGCTGC 360
QY	361 ACCTGTCAGTGGGTGACCTCTCTTGTGTCATCATCTCCCTTCTGGGCGAGTTGATGCCA 420
DB	361 ACCTGTCAGTGGGTGACCTCTCTTGTGTCATCATCTCCCTTCTGGGCGAGTTGATGCCA 420
QY	421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAAGGTGTCCATATCATCTACACTGTCA 480
DB	421 TGGCTGACTGGTACTTTGGGAAATTTTGTGTAAAGGTGTCCATATCATCTACACTGTCA 480
QY	481 ACCTCTACAGCAGGTTCTCATCTGGCCCTTATAGCTTGAACCGGTACCTCGCCATTG 540
DB	481 ACCTCTACAGCAGGTTCTCATCTGGCCCTTATAGCTTGAACCGGTACCTCGCCATTG 540
QY	541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCGAGTCTATGTGG 600
DB	541 TCCACGCCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGCGAGTCTATGTGG 600
QY	601 GCCTCTGGATCCAGCGCCCTCTCTGACTATACCTGACTTCACTTTGTCGCGAGTGCAGCC 660
DB	601 GCCTCTGGATCCAGCGCCCTCTCTGACTATACCTGACTTCACTTTGTCGCGAGTGCAGCC 660
QY	661 AGGGGGACATCAGTCAGGGGGATGA 685
DB	661 AGGGGGACATCAGTCAGGGGGATGA 685
RESULT 2	
AAZ27611	
ID	AAZ27611 standard; DNA; 1758 BP.
XX	AAZ27611;
XX	
DT	16-DEC-1999 (first entry)
XX	
DE	Mouse CXCR4 coding sequence.
DE	
KW	CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
KW	tissue repairing agent; vascularisation; ss.
XX	
OS	Mus sp.
XX	

WO9949528-A1.
 30-SEP-1999.
 23-MAR-1999; 99WO-JP001448.
 24-MAR-1998; 98JP-00095448.
 (CHUS) CHUGAI SEIYAKU KK.
 (KISH/) KISHIMOTO T.

Kishimoto T, Nagasawa T, Tachibana K;
WPI; 1999-591042/50.
P-FSDB; RAY39994.

CXCR4-potentiating agents and methods useful for inhibiting
neovascularization, and treating solid cancers.

Disclosure; Page 50-51; 63pp; Japanese.

This sequence encodes the mouse CXCR4 protein. The invention relates to
remedies inhibiting neovascularisation, remedies for solid cancer,
remedies for diseases pathologically caused by neovascularisation and
tissue repairing agents containing as the active ingredient a substance
capable of potentiating CXCR4. Based on a finding that vascularisation is
inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
remedies inhibiting vascularisation which contain as the active
ingredient a substance capable of potentiating CXCR4, remedies for solid
cancer, remedies for diseases pathologically caused by neovascularisation
and tissue repairing agents containing as the active ingredient a
substance capable of potentiating CXCR4. It is also possible to establish
methods for treatment with the use of these remedies

Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;

Query Match 82.6%; Score 566; DB 2; Length 1758;
Best Local Similarity 100.0%; Pred.No. 6.2e-156;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

120 ATGGAACCGATCAGTGAGTATATACACTCTGTGATAAACAATTCAAGAAGTCGGGTCT 179
1 ATGGAACCGATCAGTGAGTATATACACTCTGTGATAAACAATTCAAGAAGTCGGGTCT 60

180 GGAGACTATGACTCCAAAGAACCTGTTCGGGATGAAAACGTTCCAATTAAGG 239
61 GGAGACTATGACTCCAAAGAACCTGTTCGGGATGAAAACGTTCCAATTAAGG 120

240 ATTCTTCGCCACCACCATCTTCAATCATCTTTGTGACTGGCAATGCGCAAATGGAATG 299
121 ATCTTCCTGCSCCACCATCTATTCAATCATCTTCCTGACTGGCATAGTGGCAATGGAATG 180

300 GTGATCTCTGTGTATGGGTATCCAGAAGACTRAGGAGCATGACGCAACAGTACCGGCTG 359
181 GTGATCTCTGTGTATGGGTATCCAGAAGACTRAGGAGCATGACGCAACAGTACCGGCTG 240

360 CACTGTGCAGTGGTGACCTCCTCTTTGTATCACTACACTCCCCCTTCGGGCAAGTTGATGCC 419
241 CACTGTGCAGTGGTGACCTCCTCTTTGTATCACTACACTCCCCCTTCGGGCAAGTTGATGCC 300

420 ATGGCTGACTGTGACTTTGGGAAATTTTTGTGTAAAGCTGTCCATATCATCTACACTGTC 479
301 ATGGCTGACTGTGACTTTGGGAAATTTTTGTGTAAAGCTGTCCATATCATCTACACTGTC 360

480 AACCTCTACAGACGCTTCTCATCTCGCCTTCAATCAGCCTGGACCGGTACTCTGCCCAT 539
361 AACCTCTACAGACGCTTCTCATCTCGCCTTCAATCAGCCTGGACCGGTACTCTGCCCAT 420

540 GTCCACGCCACCAACAGTCAAAAGGCCAAGAACTGCTGGCTGAAAAAGCGTAGTCTATGTG 599
421 GTCCACGCCACCAACAGTCAAAAGGCCAAGAACTGCTGGCTGAAAAAGCGTAGTCTATGTG 480

600 GGCGTGTGGATGCCAGCCCTCTCTCTGACTATACCTTCAGTTTCATCTTTTGGCAGCTCAGC 659

164 TGAGGAAGTGGGCTGGAGACTATGACTCCAAAGAACCTGCTTCCGGGATGAAA 223
 127 CGAGGAATGGGCTCAGGGAGCTATGACTCCATGAAGAACCTGTTTCGTGAAGAAA 186
 224 CGTCCATTCAATAGGATCTTCTGCCACCACTACTTTCATCATCTTCTTGAATGGCAT 283
 187 TGCTAATTTCAATATAAATCTTCTGCCACCACTACTTTCATCATCTTCTTAACTGGCAT 246
 284 AGTCGCAATGGATTGGTGATCTCTGTCATGCTGTCAGTGGCTTACCAAGAACTAAGGAGCATGAC 343
 247 TGTGGGCAATGGATTGGTGATCTCTGTCATGCTGTCAGTGGCTTACCAAGAACTAAGGAGCATGAC 306
 344 GGACAAAGTACCGCTGCACCTGTGAGTGGCTGACCTCTCTTTGTTCATCAGTCCCTCTT 403
 307 GGACAAAGTACAGGCTGCACCTGTGAGTGGCTGACCTCTCTTTGTTCATCAGTCCCTCTT 366
 404 CTGGGAGTTGATGCCATGCTGACTGCTGACTGTTGGGAATTTTGTGAAGCTGTCCA 463
 367 CTGGGAGTTGATGCCATGCTGACTGCTGACTGTTGGGAATTTTGTGAAGCTGTCCA 426
 464 TATCATCTACAGCTGTCAACCTCTACAGCAGCTGTTCTCATCTCTGGCCTTTCATCAGCCTGGA 523
 427 TGTTCATCTACAGCTGTCAACCTCTACAGCAGCTGTTCTCATCTCTGGCCTTTCATCAGTCTGA 486
 524 CCGGTACCTCGCATTTGTCACCGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGA 583
 487 CCGGTACCTCGCATTTGTCACCGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGA 546
 584 AAGGCACTATGTTGGGCTGCTGGATCTCCAGCCCTCTCTCTGACTATFACCTGACTTCAT 643
 547 AAGGCACTATGTTGGGCTGCTGGATCTCCAGCCCTCTCTCTGACTATFACCTGACTTCAT 606
 644 CTTTGGCCAGCTCAGCCAGGGGAGCATCAG 673
 607 CTTTGGCCAGCTCAGTGAGGAGCATGACAG 636

RESULT 6

AAI65467
 ID AAI65467 standard; DNA; 1679 BP.
 AC AAI65467;
 XX
 DT 10-DEC-2001 (first entry)
 DE Nucleotide sequence of a human polynucleotide.
 KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
 XX C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.
 OS Homo sapiens.
 XX
 PN WO200164752-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US006322.
 XX
 PR 02-MAR-2000; 2000US-00517605.
 XX (UYN) UNIV NEW YORK STATE.
 PA (UYN-) UNIV NIJMEGEN.
 XX
 PI Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
 XX WPI; 2001-602565/68.
 DR
 XX
 PT An antibody for the treatment or prevention of HIV-infection comprises a
 FT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
 PT DC-SIGN due to concomitant conformational change.
 XX
 PS Disclosure; Page 126-127; 131pp; English.
 XX

CC The specification describes an antibody which is specific for an
 CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
 CC is exposed upon gp120 binding of DC-SIGN due to concomitant
 CC conformational change. DC-SIGN is a receptor that is specifically
 CC expressed on dendritic cells and facilitates infection of T lymphocytes
 CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
 CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
 CC with high affinity. The antibody of the invention inhibits the trans
 CC enhancement of HIV entry into a T cell or macrophage facilitated by
 CC dendritic cells. The antibody is useful to treat or prevent HIV
 CC infection. The present sequence represents a human polynucleotide, which
 CC is used in the course of the invention
 XX

SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match 65.2%; Score 446.8; DB 4; Length 1679;
 Best Local Similarity 83.0%; Pred. No. 7.7e-121;
 Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 CGAGTGCAGTACGAGTACACCTCTGAGGCGTTTGGTGTCCGGTAAACCAACCGGCTG 103
 DB 12 GCGGAGCAGGTAGCAAGTGAACCGAGGCGCTGAGTGTCCAGTAGCCACCGCATCTG 71
 QY 104 TAGACGAGTGTGGCATGGAAACGATCAGTGTGAGTATATACACTTCTGATATACTATC 163
 DB 72 GAGAACACGCGGTATACCATGGA-----GGGATCAGTATATACACTTTCAGATACTAC 126
 QY 164 TGAAGAACTGGGCTGTGAGACTATGACTCCAAAGGAAACCTGCTTCCGGGATGAAA 223
 DB 127 CGAGGAATGGCTCAGGGAGCTATGACTCCATGAAGAACCTGTTTCCGTGAAGAAA 186
 QY 224 CGTCATTTCAATAGGATCTTCTGCCACCACTACTTTCATCATCTTCTTGAATGGCAT 283
 DB 187 TGCTAATTTCAATATAAATCTTCTGCCACCACTACTTTCATCATCTTCTTAACTGGCAT 246
 QY 284 AGTCGCAATGGATTGGTGTGATCTCTGAGTGGCTTACCAAGAACTAAGGAGCATGAC 343
 DB 247 TGTGGCAATGGATTGGTGTGATCTCTGAGTGGCTTACCAAGAACTAAGGAGCATGAC 306
 QY 344 GGACAAAGTACCGGCTGCACCTGTGAGTGGCTGAGCTCTCTTTGTTCATCAGTCCCTCTT 403
 DB 307 GGACAAAGTACAGGCTGCACCTGTGAGTGGCTGAGCTCTCTTTGTTCATCAGCTTCCCTT 366
 QY 404 CTGGGCACTGATGCCATGCGGTGAGTGTGACTTGTGGGAAATTTTGTGAAGCTGTCCA 463
 DB 367 CTGGGCACTGATGCCGTTGGCAACTGTTACTTTGGGAACTTCTTATGCAAGGAGTCCA 426
 QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCTTCTCATCTCTGGCCTTTCATCAGCCTGGA 523
 DB 427 TGTTCATCTACAGCTGTCAACCTCTACAGCAGTGTCTCTATCTCTGCTGACTATFACCTGACTTCAT 643
 QY 524 CCGGTACCTCGCATTTGTCACCGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGA 583
 DB 487 CCGGTACCTCGCATTCGTCACCGCCACCAACAGTCAAGGCCAAGGAACTGTTGGCTGA 546
 QY 584 AAGGCACTATGTTGGGCTGCTGGATCTCCAGCCCTCTCTCTGACTATFACCTGACTTCAT 643
 DB 547 AAGGCACTATGTTGGGCTGCTGGATCTCCAGCCCTCTCTCTGACTATFACCTGACTTCAT 606
 QY 644 CTTTGGCCAGCTCAGCCAGGGGAGCATCAG 673
 DB 607 CTTTGGCCAGCTCAGTGAGGAGCATGACAG 636

RESULT 7

ABS53992
 ID ABS53992 standard; cDNA; 1679 BP.
 XX
 AC ABS53992;
 XX
 DT 02-DEC-2002 (first entry)
 XX
 DE DNA encoding human CXC chemokine receptor 4 (CXCR4).

XX KW CXK chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis; stroke;
KW anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW migraine; vomiting; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; dyskinesia; Huntington's disease; restenosis;
KW inflammatory bowel disease; rheumatoid arthritis; diabetic retinopathy;
KW congestive heart failure; cardiac remodeling; angiogenic diseases;
KW solid tumour; Kaposi Sarcoma; human; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 89..1147
XX FT /*tag= a
XX FT /product= "CXCR4"
XX FT /note= "CXK chemokine receptor 4"
XX
XX US2002107195-A1.
XX
XX PD 08-AUG-2002.
XX PF 17-SEP-2001; 2001US-00953692.
XX PR 21-JUL-1998; 98US-0093596P.
XX PR 21-JUL-1999; 99US-00358624.
XX PA (SMIK) SMITHLINE BEECHAM CORP.
XX PI Gupta SK;
XX WPI; 2002-697879/75.
XX P-PSDB; ABG32977.
XX
XX Inducing chemotaxis of endothelial cells, useful for regulating
XX angiogenesis, e.g. for treating cancer, comprises treatment with stromal
XX cell-derived factor 1 alpha.
XX
XX Disclosure; Fig 1; 26pp; English.
XX
XX The invention describes a method of inducing chemotaxis of endothelial
XX cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
XX method is used for stimulating EC chemotaxis, and thus angiogenesis, and
XX is used for treating atherosclerosis, restenosis, ischaemic stroke and
XX spinal cord injury. Inhibition of this process is useful in treatment and
XX prevention of a very wide range of diseases, such as, infection (by
XX bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
XX (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
XX benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
XX asthma, Parkinson's disease, acute heart failure, hypotension,
XX hypertension, urinary retention, osteoporosis, anginal pectoris,
XX myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
XX migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
XX schizophrenia) and dyskinesias (e.g. Huntington's disease), inflammatory
XX diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
XX disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
XX heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
XX and Kaposi Sarcoma. This sequence encodes the human CXK chemokine
XX receptor 4 (CXCR4) polynucleotide
XX
XX SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
XX
XX Query Match 65.2%; Score 446.8; DB 6; Length 1679;
XX Best Local Similarity 83.0%; Pred. No. 7.7e-121;
XX Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
XX
XX 44 GCAGGTGACGATGACGACCTCTCTGAGCGTTGTGCTCCGGTAACCAACCGGCTG 103

Db 12 GCGGAGCAGGTAGCAAAAGTACGCCGAGGCGCTGAGTCTCCAGTACCCCGCATCTG 71
QY 104 TAGACGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTACTC 163
Db 72 GAGAACCCAGCGGTACCATTGGA-----GGGATCAGTATATACACTTCCAGATACTACAC 126
QY 164 TGAAGAATGGGGTCTGGAGACTATGACTCCAAAGGAACCCCTCTTCCGGGTGAAA 223
Db 127 CAGGAATGGGCTCAGGGACTATGACTCCATGAAGGAACCCCTGTTTCCGTGAAGAAA 186
QY 224 CGTCATTTCAATAGGATCTTCTCCGCCACCATCTACTTCTATCTTCTTGGACTGGCAT 283
Db 187 TGCTAATTTCAATAAAATCTTCTCCGCCACCATCTACTTCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTGGTGGTCTGCTGCTATGCTGGTACCAGGAAGCTAAGGAGCATGAC 343
Db 247 TGTGGCAATGGATTGGTGGTCTGCTGCTATGCTGGTACCAGGAAGCTAAGGAGCATGAC 306
QY 344 GGACAAGTACCGGCTGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403
Db 307 GGACAAGTACCGGCTGCACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
QY 404 CTGGCAGTGTATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
Db 367 CTGGCAGTGTATGCCGCTGGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCGTCTCATCTCTGCTGCTGCTGCTGCTGCT 523
Db 427 TGTCTATCTACAGTCAACCTCTACAGCAGTGTCTCTATCTGCTGCTGCTGCTGCTGCT 486
QY 524 CCGGTACCTGCCATTTGTCACGCCACCAACAGTCAAGAGGCAAGAACTGCTGGCTGTA 583
Db 487 CCGGTACCTGCCATCTGTCACGCCACCAACAGTCAAGAGGCAAGAACTGCTGGCTGTA 546
QY 584 AAAGGCACTCTATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
Db 547 AAAGGCTGTCTATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
QY 644 CTTTCCGACGCTACGCCAGGCGGACATCAG 673
Db 607 CTTTCCCAACGTCAGTGAGGCGAGATGACAG 636
RESULT 8
ABS53750
ID ABS53750 standard; DNA; 1679 BP.
XX AC ABS53750;
XX DT 28-NOV-2002 (first entry)
XX DE DNA encoding human CXCR4 protein.
XX KW Chemotaxis; endothelial cell; EC; angiogenesis; atherosclerosis;
KW restenosis; ischaemic stroke; spinal cord injury; infection; bacteria;
KW fungi; protozoa; virus; pain; cancer; benign prostatic hypertrophy;
KW diabetes; obesity; anorexia; bulimia; asthma; allergy; hypertension;
KW Parkinson's disease; acute heart failure; hypotension; urinary retention;
KW osteoporosis; angina pectoris; myocardial infarction; stroke; dyskinesia;
KW migraine; vomiting; psychotic disorder; neurological disorder; ulcer;
KW inflammatory disorder; rheumatoid arthritis; diabetic retinopathy;
KW inflammatory bowel disease; atherosclerosis; restenosis; Kaposi sarcoma;
KW Alzheimer's disease; congestive heart failure; cardiac remodeling;
KW angiogenic disease; solid tumour; human; stromal cell derived factor-1a;
KW SDF-1alpha; CXCR4; gene; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 89..1147
XX FT /*tag= a
XX FT /product= "Human CXCR4 protein"

XX US2002107196-A1.
 XX 08-AUG-2002.
 XX 17-SEP-2001; 2001US-00953717.
 XX 21-JUL-1998; 98US-0093596P.
 XX 21-JUL-1999; 99US-00358624.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Gupta SK;
 XX WPI; 2002-706230/76.
 XX F-PSDB; ABG33065.
 XX Inducing chemotaxis of endothelial cells, useful for regulating
 XX angiogenesis, e.g. for treating cancer, comprises treatment with stromal
 XX cell-derived factor 1 alpha.
 XX Disclosure; Fig 1; 26pp; English.
 XX The present invention relates to a new method for inducing chemotaxis of
 XX endothelial cells (EC). The method of the invention involves treatment
 XX with a stromal cell-derived factor 1a (SDF1alpha). The method is used for
 XX stimulating EC chemotaxis, and thus angiogenesis, and is used for
 XX treating atherosclerosis, restenosis, ischaemic stroke and spinal cord
 XX injury, while inhibition of this process is useful in treatment and
 XX prevention of a very wide range of diseases, e.g. infection (by bacteria,
 XX fungi, protozoa or viruses), pain, cancer and benign prostatic
 XX hyper trophy, diabetes, obesity, anorexia, bulimia, asthma, allergies,
 XX Parkinson's disease, acute heart failure, hypotension, hypertension,
 XX urinary retention, osteoporosis, angina pectoris, myocardial infarction,
 XX stroke, ulcers, migraine, vomiting, psychotic and neurological disorders
 XX and dyskinesias, inflammatory disorders, rheumatoid arthritis, diabetic
 XX retinopathy, inflammatory bowel disease, atherosclerosis, restenosis,
 XX Alzheimer's disease, congestive heart failure, cardiac remodeling,
 XX angiogenic diseases, solid tumours, and Kaposi sarcoma. The present
 XX nucleic acid sequence encodes the human CXCR4 protein of the invention.
 XX Note: This sequence represents SEQ ID NO 1 given in figure 1 of the
 XX specification

XX SQ Sequence 1679 BP; 407 A; 399 C; 367 G; 504 T; 0 U; 2 Other;
 Query Match 65.2%; Score 446.8; DB 6; Length 1679;
 Best Local Similarity 83.0%; Pred. No. 7.7e-121;
 Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
 44 GCAGGTGACAGTACAGTACCCCTCTGAGCGCTTTGGTCTCGGTACACACACGCTG 103
 12 GCGGCAGCAGGTACGAAAGTGACGCCGCGGCTGAGTCTCCAGTAGCCCGCATCTG 71
 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTATCTC 163
 72 GAGAACGACGGTTACATGGA-----GGGATCAGTATATACACTTCAATCACTACAC 126
 164 TGAAGAAGTGGGTCTGGAGTATGACTCCAAAGGAACCGTCTCCGGATGA AAA 223
 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCGTCTCCGTGAAGAAA 186
 224 CTTCCATTTCAATAGATCTTCTGCGCCACCTACTCATCATCTTCTGACTGGCAT 283
 187 TGTAAATTTCAATAAATCTTCTGCGCCACCTACTCATCATCTTCTTAACTGGCAT 246
 284 AGTCGGAATGGATTTGGTATCTCTGCTGATGGTTTACAGAGAAGCTAAGGAGCATGAC 343
 247 TGTGGCAATGGATTTGGTATCTCTGCTGATGGTTTACAGAGAAGCTAAGGAGCATGAC 306
 344 GGAACAGTACCGGTGACCTGTCTGAGTGGTACCTCTCTTTGTATCATCACATCCCTTT 403
 307 GGAACAGTACAGCTGACCTGTCTGAGTGGCGCCACCTCTCTTTGTATCATCACGCTTCT 366

QY 404 CTGGGCAGTTGATGCCATGGCTGACTGGTACTTTTGGAAATTTTCTGTAAGGCTGTCCA 463
 Db 367 CTGGGCAGTTGATGCCGTCGCAAACTGGTACTTTTGGAACTTCTATGCAAGCAGTCCA 426
 QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCGTTCTCATCTCTGCGCTTATCATGCGTGA 523
 Db 427 TGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTCATCTCTGCGCTTATCATGCTGGA 486
 QY 524 CCGGTACTCGCATTTGTCCACGCCACCACTCAAGGCCCAAGAACTGCTGGCTGA 583
 Db 487 CCGTACTCTGCGCATGTCCACGCCCACTCAAGGCCCAAGAACTGCTGGCTGA 546
 QY 584 AAAGGCACTCTATGTGGCGTCTGGATCCAGCCCTCTCTGACTATATACCTGACTTCAT 643
 Db 547 AAAGTGTCTATGTGGCGTCTGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
 QY 644 CTTTCCGACGTGACGAGGGGACATCAG 673
 Db 607 CTTTCCCAACGTGACGTGAGGACATGACAG 636
 RESULT 9
 ABV78039
 ID ABV78039 standard; DNA; 1679 BP.
 XX AC ABV78039;
 XX DT 12-NOV-2002 (first entry)
 XX DE Hypoxia-regulated protein coding sequence #59.
 XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 XX antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
 XX hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 XX ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 XX preclapnea; atherosclerosis; inflammatory condition; wound healing;
 XX inflammation; erythropoiesis; hair loss; human; gene; ds.
 XX Homo sapiens.
 OS WO200246465-A2.
 PN 13-JUN-2002.
 PD 10-DEC-2001; 2001WO-GB005458.
 PF 08-DEC-2000; 2000GB-00030076.
 PR 08-FEB-2001; 2001GB-00003156.
 PR 25-OCT-2001; 2001GB-00025666.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
 PI Rayner WN;
 XX WPI; 2002-627238/67.
 XX Identifying a gene involved in disease for treating hypoxia-regulated
 XX conditions, comprises comparing the transcriptome/proteome of two cell
 XX types under different conditions and identifying a differentially
 XX regulated gene.
 XX Claim 37; Page 378; 538pp; English.
 XX The present invention relates to methods for identifying genes and
 XX proteins that are implicated in a specific disease or physiological
 XX condition. The method comprises comparing the transcriptome/proteome of a
 XX specialised cell type implicated in a disease or condition with that of a
 XX second specialised cell type, under two experimental conditions, and
 XX identifying a gene that is differentially regulated in the two
 XX specialised cell types under experimental conditions. ABV77873-ABV78116
 XX and ABV65061-ABV65257 were identified using the methods of the invention.
 XX The coding sequences and proteins are useful for treating a disease in a

CC patient, for manufacture of a medicament for treating hypoxia-regulated
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
CC biological response to hypoxia conditions, or hypoxic-associated
CC pathology in a patient. The coding sequences and proteins are also useful
CC for monitoring the therapeutic treatment of a disease or physiological
CC condition, such as cancer, ischaemic conditions, reperfusion injury,
CC retinopathy, neonatal stress, pre-eclampsia, atherosclerosis, inflammatory
CC conditions, wound healing, inflammation, erythroptosis or hair loss
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match	65.2%	Score 446.8;	DB 6;	Length 1679;
Best Local Similarity	83.0%;	Pred. No. 7.7e-121;		
Matches 523;	Conservative	0;	Mismatches 102;	Indels 5;
				Gaps 1;

QY	44	CGAGTGCAGTAGCAGTGACCCCTCTCAGCGGTTTGGTGTCTCCGGTAAACACCAACGGCTG	103
Db	12	CGCGCAGCAGTAGCAAGTAGACGCCAGGCGCTGAGTGTCTCAGTAGTCAACCGCATCTG	71
QY	104	TAGACCGAGTGTTCGCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATAAATACTCTC	163
Db	72	GAGAACCAACGCGTTTACCATGCA-----GGGATTCAGTTATATACACTTCAGATAACTACAC	126
QY	164	TGAGAAAGTGGGCTCGGAGACTATGACTCCACACAGGACCCGTCTCGGGATGAAAA	223
Db	127	CGAGGAATAGGGCTCAGGGHACTATGACTCATGAAGAACCTGTTTCCGTGAAGAAAA	186
QY	224	CGTCAATTTCAATAGGATCTTCTCGCCCAACATCTACTTCATCATCTTCTTGACTGGCAT	283
Db	187	TGCTAATTTCAATAAAATCTTCTCGCCCAACATCTACTCCATCATCTTCTTAATGGCAT	246
QY	284	AGTCGGCAATAGATTGGTGTGATCTCTGTCATGGGTACCAGAAAGCTAAGGAGCATGAC	343
Db	247	TGTGGGCAATGGATTGGTGCATCTTGTCATCTGGGTTACCAGAAAGAACTGAGAAGCATGAC	306
QY	344	GGACAAGTACCGGCTGCACCTGTTCAGTGGGTGAACCTCTTTGTTCATCACACTCCCTTT	403
Db	307	GGACAAGTACAGGGCTGCACCTGTTCAGTGGCGCACTCTCTTTGTTCATCACGCTTCCTTT	366
QY	404	CTGGCAGTTGATGTCATCGGCTGACTGTGTACTTTGGGAAATTTTGTGTAAAGCTGTCCA	463
Db	367	CTGGCAGTTGATGCGCTGGCAAACTGGTACTTTGGGAACTTCTCATGCAAGGCAGTCCA	426
QY	464	TATCATCTACTGTCAAACTCTTCAGAGCGTTTCTCATCTCGGCTTCAATCAGCGCTGGA	523
Db	427	TGTCATCTACACAGTCAAACTCTACAGCAGTGTCTCATCTCGGCTTTCATCAGTCTGGA	486
QY	524	CCGCTACCTTCGGCATTTGTCCACGGCCACCACACAGTCAAAAGGCCAAGGAAACTGTCTGGCTGA	583
Db	487	CCGCTACCTTCGGCCATCGTCCACGCCCAACACAGTCAGAGGCCAAGGAAGCTGTTGGCTGA	546
QY	584	AAAGGAGTCTATGTGGCGGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT	643
Db	547	AAAGGTGGTCTATGTTGGCGTCTGGATCCCTGCGCTCTCTGACTATTCCTCGACTTCAT	606
QY	644	CTTTGCCACAGTTCAGCCAGGGGGACATCAG	673
Db	607	CTTTGCCACAGTTCAGTGAAGGCAGATGACAG	636

RESULT 10

RES00110
 AEX08779
 ID AEX08779 standard; cDNA; 1679 bp.
 XX
 AC AEX08779;
 XX
 XX
 DT 21-JAN-2003 (first entry)
 XX
 DE Angiogenesis-associated human polynucleotide sequence #41.
 XX
 KW Human; angiogenesis-associated transcript; angiogenesis;
 KW angiogenesis-associated disease; cancer; cytostatic; gene therapy; gene
 KW ss.

XX	Homo sapiens.
OS	
XX	
XX	WO200279492-A2.
PN	
XX	
XX	10-OCT-2002.
PD	
XX	
XX	14-FEB-2002; 2002WO-US004915.
PF	
XX	
XX	14-FEB-2001; 2001US-00784356.
PR	
PR	22-FEB-2001; 2001US-00791390.
PR	19-APR-2001; 2001US-0285475P.
PR	03-AUG-2001; 2001US-0310025P.
PR	13-NOV-2001; 2001US-0350666P.
PR	29-NOV-2001; 2001US-0334244P.
XX	
XX	(EOSB-) EOS BIOTECHNOLOGY INC.
PA	
XX	
XX	Murray R, Glynne R, Watson SR, Aziz N;
PI	
XX	
XX	WPI; 2003-040681/03.
DR	
DR	P-PSDB; ABU03495.
XX	
XX	Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT	treating cancer by contacting a sample with a polynucleotide that
PT	exhibits changes in expression level as a function of time in tissue
PT	undergoing angiogenesis.
PT	

Example 2: Page 217-218: 291pp: English.

The present invention relates to methods and compositions for detecting an angiogenesis-associated transcript in a cell in a patient. The method involves contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence at least 80% identical to any of the angiogenesis-associated human polynucleotide sequences given in the specification. These angiogenesis-associated polynucleotide sequences comprise genes that exhibit changes in expression levels as a function of time in tissue undergoing angiogenesis. The method and the polynucleotide sequences of the invention are useful for diagnosing and treating angiogenesis and angiogenesis-associated diseases e.g. cancer. The polynucleotide sequences are also useful in the gene therapy of such disorders. The angiogenesis-associated proteins encoded by the polynucleotide sequences are useful as a vaccine for therapeutic and prophylactic immunisation. ABX08739-ABX08853 represent angiogenesis-associated polynucleotide sequences

Sequence 1679 BP: 407 A: 400 C: 368 G: 504 T: 0 U: 0 Other:

Query Match 65.2%; Score 446.8; DB 7; Length 1679;
Best Local Similarity 83.0%; Pred. No. 7.7e-121;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

44	Qy	GCAGGTGCAGGTAGCAGTGCACCCCTCTGAGCGGTTTGGTGTCTCGGTAAACACACCGGCTG	103
12	Db	CGCGCAGCAGGTAGCAAAAGTGACCGCGAGGCCCTGAGTGTCTCCAGTAGCGCACCGCATCTG	71
104	Qy	TAGAGCGAGTGTTCCTCCATCGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTC	163
72	Db	GAGAACGAGCGGTTACCATGGA-----GGGGATCAGTATATATACACTTCAGATAACTACAC	126
164	Qy	TGAAGAAGTGGGGTCTGGAGACATATGACTCCAAACAGGAACCTCGTTCCGGGATGAAAA	223
127	Db	CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTCCGTGAAGSAAA	186
224	Qy	CGTCCATTTCAATAGAGATCTTCTCTGCCACCACTACTTCATCATCTTCTTTGACTGGCAT	283
187	Db	TGCTAATTTCAATAAAAATCTTCTCTGCCCACTACTCTACTCCATCATCTTCTTAATGGCAT	246
284	Qy	AGTCGCAATGGATTGGTGATCTGTGTCTATGGGTTACCAAGAAAGCTAAGGAGCATGAC	343
247	Db	TGTGGCCAAATGGATTGGTGATCTGTGTCTATGGGTTACCAAGAAAGCTAAGGAGCATGAC	306

QY 344 GGACAGTACCGCTGCACTGTGAGTGGGTGACCTCTCTTTGTTCATCACATCCCTTT 403
 DB 307 GGACAAAGTACAGGCTGACCTGTGAGTGGCGGACCTCTCTTTGTTCATCACGCTTCCCTT 366
 QY 404 CTGGGAGTTGATGGCACTGGCTGACTGCTTGTGGAAATTTTGTGTAAGCTGTCCA 463
 DB 367 CTGGGAGTTGATGGCGTGGCAAACTGGTACTTTGGAACTTCTTATGAGGCGATCCA 426
 QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGCGTTCTCATCTGGCCCTTCATCAGCCTGGA 523
 DB 427 TGTCACTACAGTCAACCTCTACAGCAGTGTCTCATCTGGCCCTTCATCAGTCTGGA 486
 QY 524 CCGGTACCTCGCCATGTTCACGCCCAACCAAGTCAAGGCCAAGAACTGCTGCTGA 583
 DB 487 CCCTACCTGGCCATGTTCACGCCCAACCAAGTCAAGGCCAAGAACTGCTGCTGA 546
 QY 584 AAAGGAGTCTATGTGGGGCTCTGGATCCCGAGCCCTCTCTGACTATACCTGACTTCAT 643
 DB 547 AAAGGTGGTCTATGTGGGGCTCTGGATCCCGAGCCCTCTCTGACTATACCTGACTTCAT 606
 QY 644 CTTTGGCAGCTCAGCCAGGGGAGATCAG 673
 DB 607 CTTTGGCAGCTCAGTGAGGAGATGACAG 636

RESULT 11

ABX74454

ID ABX74454 standard; cDNA; 1679 BP.

XX

AC ABX74454;

XX 21-MAR-2003 (first entry)

XX

XX Human cDNA sequence #26 up-regulated in CC-RCC patients.

XX

XX Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.

XX

OS Homo sapiens.

XX

XX WO200279411-A2.

XX

XX 10-OCT-2002.

XX

XX 29-MAR-2002; 2002WO-US009576.

XX

XX 29-MAR-2001; 2001US-0279411P.

XX

XX (VAND-) VAN ANDEL INST.

XX

XX Haab B, Rhodes D, Teh BT, Takashi M;

XX

XX WPI; 2003-040679/03.

XX

XX New microarray, comprising a matrix of cDNA probe from a set of probes
 PT immobilized to a solid surface in predetermined order, useful in the
 PT prognosis of patients with clear cell renal carcinoma.

XX

XX Claim 35; Page 132-133; 179pp; English.

XX

XX The present invention relates to a microarray comprising a matrix of at
 CC least one cDNA probe from a set of probes immobilised to a solid surface
 CC in a predetermined order, where a row of pixels corresponds to replicates
 CC of one distinct probe from the set. The probes are complementary to
 CC nucleic acid sequences that are expressed differentially in aggressive as
 CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
 CC and that hybridise to the probes under high stringency conditions. The
 CC microarray is useful for the prognosis of patients with CC-RCC, wherein
 CC aggressive and non-aggressive CC-RCC tumour types are characterised by
 CC differential expression profiles of genes that hybridise with one or more
 CC probes immobilised on the microarray. The arrays are useful for gene

CC expression profiling of tumour and normal tissues. The present sequence
 CC represents a human cDNA sequence up-regulated in CC-RCC patients
 XX
 SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match

Best Local Similarity 65.2%; Score 446.8; DB 7; Length 1679;

Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGAGGTAGCAGTACCTCTGAGGCGTTTGGTCTCCGGTAAACCAACCAACGGCTG 103

DB 12 GGGCAGCAGGTAGCAAGTACGCGCGAGGCGCTGAGTCTCCAGTAGCCACCGCATCTG 71

QY 104 TAGACGAGTGTGCCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATATACTACTC 163

DB 72 GAGAACCCAGCGGTTTACCATGGA-----GGGGATCAGTATATATACACTTCTGATATACTACTC 126

QY 164 TGAAGAGTGGGGCTGAGAGATATGACTCTCAACAGAGAAACCTGCTTCGGGGATGAAA 223

DB 127 CGAGAAATGGGCTCAGGGGACTAAGTCTCATGAGAGAAACCTGTTTCCGTGAAGAAA 186

QY 224 CGTCCATTTCAATAGGATCTTCTGCCCACCATCTACTTTCATCATCTTCTTGTGACTGSCAT 283

DB 187 TGCTAAATTTCAATAAATCTTCTGCCCACCATCTACTTTCATCATCTTCTTAACTGGCAT 246

QY 284 AGTCGGCAATGGATGGTCTGATCCTGGTTCATGGTTTACAGAGAGCTAAGGACATGAC 343

DB 247 TGTGGCAATGGATGGTCTGATCCTGGTTCATGGTTTACAGAGAGAACTGAGAAAGCATGAC 306

QY 344 GGACAAAGTACCGGCTGCACCTGTGAGTGGGTGACCTCTCTTTGTTCATCACACTTCCCTTT 403

DB 307 GGACAAAGTACAGGCTGCACCTGTGAGTGGCGGACCTCTCTTTGTTCATCACGCTTCCCTT 366

QY 404 CTGGGAGTTGATGCCATGGTGTGACTGTTGGGAAATTTTGTGTAAGGCTGTCCA 463

DB 367 CTGGGAGTTGATGCCGTGGCAACTGGTACTTTTGGGAACTTCTATGCAAGGCGAGTCCA 426

QY 464 TATCATCTACACTGTCAACCTCTACAGCAGGTTTCTCATCTGCGCTTCTATCAGCCTGGA 523

DB 427 TGTCACTACACAGTCAACCTCTACAGCAGTGTCTCATCTGCGCTTCTATCAGTCTGGA 486

QY 524 CCGTACTCTGCCATTTGTCCAAGCCACCAACAGTCAAAAGGCCAAGAAACTGTGCGTGA 583

DB 487 CCGTACTCTGCCATCTGTCCACGCCCAACAGTCAAGGCCAAGAAAGCTGTGCGTGA 546

QY 584 AAAGGAGTCTATGTGGGGCTGAGTCCCGAGCCCTCTCTGACTATACCTGACTTTCAT 643

DB 547 AAAGGTGCTATGTGGGGCTGAGTCCCTGCGCTCTCTGACTATACCTGACTTTCAT 606

QY 644 CTTTGGCAGCTCAGCCAGGGGAGATCAG 673

DB 607 CTTTGGCAGCTCAGTGAGGAGATGACAG 636

RESULT 12

ABZ68886

ID ABZ68886 standard; cDNA; 1679 BP.

XX

XX ABZ68886;

XX

XX 28-MAY-2003 (first entry)

XX

XX Nucleotide sequence of human chemokine receptor CXCR4.

XX

XX Human; chemokine receptor; CXCR4; viral infection; surface protein;
 KW respiratory virus infection; respiratory syncytial virus infection;
 KW RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX

XX 89..1147

XX /*tag= a

FT XX /product= "CXCR4"
FN WO2003014153-A2.
XX
PD 20-FEB-2003.
XX
PF 12-AUG-2002; 2002WO-CA001248.
XX
PR 10-AUG-2001; 2001US-0311088P.
XX
XX (TOPI-) TOPIGEN PHARM INC.
PA
XX Renzi P, Zemzoui K;
PI WPI; 2003-256541/25.
XX P-PSDB; ABP97733.
DR
DR
XX
XX
PT Modulating viral infection of a cell, for treating or preventing
PT respiratory virus infections, bronchitis, pneumonia or asthma, by
PT modulating a binding interaction between a cell chemokine-receptor and a
PT surface protein of the virus.
XX
XX
PS Disclosure; Page 117-119; 120pp; English.
XX
XX The present sequence encodes human chemokine receptor CXCR4. The
CC specification describes a method for modulating viral infection of a
CC cell. The method comprises modulating a binding interaction between a
CC cell chemokine-receptor and a surface protein of the virus. The proviso
CC is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC not HIV. The method is useful for treating or preventing respiratory
CC virus infection in vertebrates, more particularly respiratory syncytial
CC virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC bronchitis, pneumonia or asthma
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 65.2%; Score 446.8; DB 7; Length 1679;
Best Local Similarity 83.0%; Pred. No. 7, 7e-121;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGTGCAGGTAGCAGTACCCCTCTGAGGCGTTTGGTGGCTCCGGTAACCAACCGGCTG 103
DB 12 GCGGCAGCAGGTAGCAAAAGTGAAGCGCGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
QY 104 TAGAGCGAGTGTGGCATGGAACCGATCAGTGTGAGTATATACACTTCTGTAACTACTC 163
DB 72 GAGAACCGCGTTTACCATGGA-----GGGATCAGTATATACACTTCAGTAACTACAC 126
QY 164 TGAAGAAGTGGGGTCTGGAGACTATGACTTCCAAAGGAACCTCTCTCCGGGATGAAAA 223
DB 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCTCTTTCGGTGAAGAAA 186
QY 224 CTTCCATTTTCATAGGATCTTCCGCGCAGCATCTACTTCTCTTCTGACTGGCAT 283
DB 187 TCTTAATTTCAATAAAATCTTCTCCGCCACCATCTACTCCATCATCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGATGTGGTGTATCTCTGGTCTATGGGTTTACCAGGAAGCTAAGGAGCATGAC 343
DB 247 TGTGGCAATGATGTTGGTCACTCTCTGTATGGTTTACCAGGAAGAACTGAGAAGCATGAC 306
QY 344 GGAAGTACCGGCTGACCTGTGAGTGGTGACCTCTCTTGTTCATCAGTCCGCTT 403
DB 307 GGACAAGTACAGGCTGCACTGTGAGTGGCGGACCTCTCTTGTGTCATGAGCTTCCCTT 366
QY 404 CTGGGCGAGTTGATGCCATGGCTGACTGTCTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
DB 367 CTGGGCGAGTTGATGCGGTGGCAAACTGTGTTGGGAATCTCTATGCAAGGAGTCCA 426
QY 464 TATCATCTACAGTGCACCTCTAGCAGAGGTTCTCATCTCTGGCTTTCATCAGCTGGA 523
DB 427 TGTCTATCTACAGTCAACCTCTACAGCAGTGTCTCTATCTCTGGCTTTCATCAGTCTGGA 486
QY 524 CCGGTACCTCGCCATTGTCTCCAGCCACCAACAGTCAAGGCCAAGGAACTGCTGGCTGA 583

DB 487 CCGTACCTGGCCATCGTCCACGACCAACAGTCAGAGGCCAAGGAGCTGTGGCTGA 546
QY 584 AAAGCAGTCTATGTGGGCTGTGGATCCAGCCCTCTCTGACTATACCTGACTTCAAT 643
DB 547 AAAGGTGGTCTATGTGGGCTGTGGATCCCTCTCTGACTATTCCTCCGACTTCAAT 606
QY 644 CTTTGGCCAGCTCAGCCAGGCGGACATCAG 673
DB 607 CTTTGGCCAAAGTCAAGTGAGGCGAGATGACAG 636
RESULT 13
AB242642
ID AB242642 standard; DNA; 1679 BP.
XX
AC AB242642;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human CXCR chemokine receptor 4 nucleotide SEQ ID NO:75.
XX
XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor; modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
XX 19-DEC-2000; 2000US-0257144P.
XX
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
XX Burmer GC, Roush CL, Brown JP;
PI WPI; 2003-046718/04.
DR P-PSDB; ABP81796.
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
XX Disclosure; Fig 1; 523pp; English.
XX
XX The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,

CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match 65.2%; Score 446.8; DB 7; Length 1679;
Best Local Similarity 83.0%; Pred. No. 7,7e-121;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGCGAGTACAGTACCTCTGAGCGGTTTGGTCTCGGTAAACACACCGGCTG 103
Db 12 GCGGACGAGGTAGCAAGTACGCGGAGGCGCTGAGTCTCCAGTAGCCACCGCATCTG 71

QY 104 TAGAGCGAGTGTGCGATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTCTC 163
Db 72 GAGAACCGCGGTACCATGGA-----GGGGATCAGTATATACACTTCAGTAACTACAC 126

QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCAAAGGAACCTGCTCCGGGATGAMAA 223
Db 127 CGAGGAATGGGCTCAGGAGCTATGCTCCAGAGAACCTGTTTCGTTGAGAGAA 186

QY 224 CGTCCATTTCAATAGGATCTTCTGCGCCACCACTACTCATCTTTGAGTGGCAT 283
Db 187 TGCTAATTTCAATAAATCTTCTGCGCCACCACTACTCATCTTTTAACTGGCAT 246

QY 284 AGTCGCAATGGATGTGTGATCTGCTGATGCTTACAGAGAACTAAGGAGCATGAC 343
Db 247 TGTGGCAATGGATGTGTGATCTGCTGATGCTTACAGAGAACTGAGAGCATGAC 306

QY 344 GGCAAGTACCGGTGACCTGTGAGTGGTGACCTCTCTTTGTATCAGTACCTCCCTT 403
Db 307 GGCAAGTACAGGTGACCTGTGAGTGGTGACCTCTCTTTGTATCAGTACCTCCCTT 366

QY 404 CTGGGAGTGTGATGCGGCTGAGTGTGATGCTTGGGAAATTTTGTAAAGCTGTCCA 463
Db 367 CTGGGAGTGTGATGCGGCTGAGTGTGATGCTTGGGAAATTTTGTAAAGCTGTCCA 426

QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCTTCTCATCTCGGCTTTCATCAGCTGGA 523
Db 427 TGTCACTACAGTCAACCTCTACAGCAGTCTCTCATCTCGGCTTTCATCAGCTGGA 486

QY 524 CGGTACTCGCATTTGTTCAGCGCCACACAGTCAAGGCCAAGGAATCTGCTGCTGA 583
Db 487 CGGTACTCGCATTTGTTCAGCGCCACACAGTCAAGGCCAAGGAATCTGCTGCTGA 546

QY 584 AAGGCGAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT 643
Db 547 AAGGCGAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTCAT 606

QY 644 CTTTGGCGAGCTCAGCCAGGGGAGATCAG 673
Db 607 CTTTGGCGAGCTCAGTGGGAGATGACAG 636

RESULT 14
ID ADC98645
XX ADC98645 standard; cDNA; 1679 BP.
AC
AC ADC98645;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human CXC chemokine receptor 4 (CXCR4)-encoding cDNA.
DE
XX Human; CXC chemokine receptor 4; CXCR4; CXCR4 inhibitor;
KW

KW small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
KW metastasis; motility; morphological change; drug screening; monitoring;
KW cytostatic; gene; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 89..1147
FT /*tag= a
FT /product= "Human CXCR4"
XX
XX WO2003079020-A2.
XX
XX 25-SEP-2003.
XX
XX 20-MAR-2003; 2003WO-EP002916.
XX
XX 20-MAR-2002; 2002US-0366370P.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX (NOVS) NOVARTIS AG.
XX Salgia R;
XX
XX WPI; 2003-757037/71.
XX P-PSDB; ABC98646.
XX
XX Use of a CXC chemokine receptor 4 (CXCR4) inhibitor, for modulating
XX cellular adhesion or inhibiting proliferation, movement, or morphological
XX change in a small cell lung cancer (SCLC) cell population, or for
XX treating a subject having SCLC.
XX
XX Disclosure; SEQ ID NO 1; 78pp; English.

XX The invention relates to methods and compositions for inhibiting small
XX cell lung cancer (SCLC) proliferation and metastasis through modulation
XX of the activity or expression of CXC chemokine receptor 4 (CXCR4), and
XX optionally, that of the tyrosine kinase receptor c-Kit. Administration of
XX a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
XX proliferation, motility, and morphological changes in an SCLC cell
XX population. The invention is based on the finding that CXCR4 is
XX ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
XX The invention also discloses methods of screening for CXCR4 inhibitors;
XX determining whether a sample of lung cancer cells expresses CXCR4 and is
XX therefore amenable to CXCR4 inhibitor therapy; determining whether a
XX sample of lung cancer cells coexpress CXCR4 and c-Kit is is therefore
XX amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
XX inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
XX CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
XX by determining CXCR4 (and optionally c-Kit) activity in two or more
XX tumour samples taken from an SCLC patient during the course of treatment.
XX The methods of the invention are useful for treating SCLC, for screening
XX for agents for the treatment of SCLC, for determining whether a SCLC from
XX individual patients is suitable for treatment with agents of the
XX invention, and for determining whether such treatment is effective and
XX should be continued. The present sequence represents cDNA encoding human
XX CXCR4, which can be used in methods of the invention.

XX Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
SQ

Query Match 65.2%; Score 446.8; DB 9; Length 1679;
Best Local Similarity 83.0%; Pred. No. 7,7e-121;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGCGAGTACAGTACCTCTGAGCGGTTTGGTCTCGGTAAACACACCGGCTG 103
Db 12 GCGGACGAGGTAGCAAGTACGCGGAGGCGCTGAGTCTCCAGTAGCCACCGCATCTG 71

QY 104 TAGAGCGAGTGTGCGATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTCTC 163
Db 72 GAGAACCGCGGTACCATGGA-----GGGGATCAGTATATACACTTCAGTAACTACAC 126

QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCAAAGGAACCTGCTCCGGGATGAMAA 223

Db 642 CTTTGCCAAAGTCAGTCAGTGAGGCGACATGACAG 671

Search completed: May 17, 2004, 06:37:07
Job time : 295.749 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 54.8443 Seconds
(without alignments)
6931.281 Million cell updates/sec

Title: US-09-367-052-5

Perfect score: 685

Sequence: 1 ccacctaatacagactact.....gacatcagtcgggggatga 685

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS-COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	446.8	65.2	1679	4	US-09-517-605-14
2	445.8	65.1	1664	4	US-09-582-224A-5
3	445.8	65.1	1664	4	US-09-023-655-1213
4	439.2	64.1	1737	1	US-08-202-056-4
5	439.2	64.1	1737	1	US-08-076-093A-3
6	439.2	64.1	1737	1	US-08-701-265-3
7	439.2	64.1	1737	2	US-08-284-586-3
8	439.2	64.1	1737	2	US-08-805-478-3
9	439.2	64.1	1737	2	US-08-802-627A-3
10	439.2	64.1	1737	2	US-08-801-238-3
11	439.2	64.1	1737	2	US-08-801-228-3
12	439.2	64.1	1737	3	US-09-104-296-3
13	439.2	64.1	1737	5	PCT-US94-06380-2
14	430.8	62.9	1225	4	US-09-016-434-1235
15	428	62.5	1317	1	US-08-153-848-45
16	428	62.5	1317	3	US-09-299-843A-45
17	428	62.5	1317	4	US-09-088-337B-45
18	428	62.5	1317	5	PCT-US93-11153-45
19	152.4	22.2	1119	4	US-09-170-496D-65
20	152.4	22.2	1119	4	US-09-170-496D-199
21	152.4	22.2	1679	1	US-08-202-056-6
22	152.4	22.2	1679	1	US-08-076-093A-5
23	152.4	22.2	1679	1	US-08-701-265-5
24	152.4	22.2	1679	2	US-08-284-586-5
25	152.4	22.2	1679	2	US-08-805-478-5
26	152.4	22.2	1679	2	US-08-802-627A-5
27	152.4	22.2	1679	2	US-08-801-238-5

28	152.4	22.2	1679	2	US-08-801-228-5	Sequence 5, Appli
29	152.4	22.2	1679	3	US-09-104-296-5	Sequence 5, Appli
30	152.4	22.2	1679	5	PCT-US94-06380-3	Sequence 3, Appli
31	152.4	22.2	2818	3	US-08-992-493-7	Sequence 7, Appli
32	152.4	22.2	2818	3	US-08-628-655-1	Sequence 1, Appli
33	141.2	20.6	2085	3	US-09-299-843A-65	Sequence 65, Appli
34	141.2	20.6	2085	4	US-09-088-337B-65	Sequence 65, Appli
35	139.4	20.4	1029	3	US-09-116-498-5	Sequence 5, Appli
36	138.8	20.3	2751	1	US-08-153-848-23	Sequence 23, Appli
37	138.8	20.3	2751	3	US-09-299-843A-23	Sequence 23, Appli
38	138.8	20.3	2751	4	US-09-088-337B-23	Sequence 23, Appli
39	138.8	20.3	2751	5	PCT-US93-11153-23	Sequence 23, Appli
40	137.8	20.1	1037	3	US-09-116-498-3	Sequence 3, Appli
41	131.8	19.2	1107	4	US-09-170-496D-19	Sequence 19, Appli
42	131.8	19.2	1107	4	US-09-170-496D-173	Sequence 173, App
43	131.8	19.2	1293	4	US-09-016-434-1052	Sequence 1052, Ap
44	131.8	19.2	1670	3	US-08-709-838-1	Sequence 1, Appli
45	131.8	19.2	1670	3	US-08-829-839-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-517-605-14
; Sequence 14, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-517-605-14

Query Match					65.2%; Score 446.8; DB 4; Length 1679;
Best Local Similarity					83.0%; Pred. No. 1e-116;
Matches 523; Conservative					0; Mismatches 102; Indels 5; Gaps 1;
Qy	44	GCAGTCAGGTAGCAGTGACCTCTGAGCGGTTGGTGCTCCGGTAACCCACCGCTG	103		
Db	12	CGGCACGCGGTAGCAAGTGACGCGGCGCTGAGTGCTCCAGTAGCCACCGCATCTG	71		
Qy	104	TAGAGCGAGTGTGCCCATGGAACCGATCAGTGTGAGTATATACACTTCTTGATACTACTC	163		
Db	72	GAGAACCGCGGTACCATGGA-----GGGATCAGTATATACACTTCAGATACTACAC	126		
Qy	164	TGAAGAGTGGGTCTGGAGACTATGACTCAACAGGAACCTGCTTCGGGATGAAA	223		
Db	127	CGAGGAAATGGGCTCAGGGAGCATGACTCCATGAAGAACCCCTGTTCCGTGAAGAAA	186		
Qy	224	CGTCCATTCAATAGGATCTTCCTGCCACCATCTACTTCATCATCTTCTTGACTGGCAT	283		
Db	187	TGCTAATTTCAATAAATCTTCCTGCCACCATCTACTTCATCATCTTCTTAATGGCAT	246		
Qy	284	AGTCGGCAATGGATGTGATCTCTGGTCACTGGGTACGAAGAGTATAGGAGCATGAC	343		
Db	247	TGTGGCAATGGATGTGATCTCTGGTCACTGGGTACGAAGAACTGAGAAGCATGAC	306		
Qy	344	GGACAGTACGGCTGCACCTGTCAGTGGTGGCTCTCTTTGTATCATCATCTCCCTT	403		
Db	307	GGACAGTACGGCTGCACCTGTCAGTGGCTGGCTCTCTTTGTATCATCATCTCCCTT	366		

QY 404 CTGGGAGTTGATGCCATGGCTGACTGGTACTTTTGGGAAATTTTGTGTAAGGCTGTCCA 463
Db 367 CTGGGAGTTGATGCCATGGCTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGCACTCCA 426
QY 464 TATCATCTACAGTCTCAACCTCTACAGACGGTTCTCATCCTGGCCTTCATCAGCCTGGA 523
Db 427 TGTCTATACAGTCTCAACCTCTACAGACGGTTCTCATCCTGGCCTTCATCAGTCTGGA 486
QY 524 CCGTACTCTGGCCTTCTCCAGCCCAACAGTCAAAAGGCCAAGAACTGCTGGCTGA 583
Db 487 CCGTACTCTGGCCTTCTCCAGCCCAACAGTCAAAAGGCCAAGAACTGCTGGCTGA 546
QY 584 AAAGGAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCAT 643
Db 547 AAAGTGGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATTCCTGACTTCAT 606
QY 644 CTTTGGCGAGTCTAGCCAGGGGAGCATCAG 673
Db 607 CTTTGGCAAGTCTAGTGAGGAGATGACAG 636

RESULT 2
US-09-582-224A-5
; Sequence 5, Application US/0958224A
; Patent No. 6429308
; GENERAL INFORMATION:
; APPLICANT: Iijima, Takeshi
; APPLICANT: GOTO, Takeshi
; APPLICANT: SHIMADA, Takashi
; TITLE OF INVENTION: HIV Infection Inhibitors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/582,224A
; PRIORITY FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: PCT/JP99/06534
; PRIORITY FILING DATE: 1999/11/24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match 65.1%; Score 445.8; DB 4; Length 1664;
Best Local Similarity 83.0%; Pred. No. 1.9e-116;
Matches 522; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 45 CAGGTGAGGTAGCAGTACCCCTCTGAGGCGTTTGGTCCCGTAAACCACCAACGCGTGT 104
Db 1 CGGCAGCAGGTAGCAAGTGAAGCGGAGGCGCTGAGTCCAGTAGCCACCGCATCTGG 60
QY 105 AGAGGAGTGTGGCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATTAATACTCT 164
Db 61 AGAACCCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCCAGATAACTACCC 115
QY 165 GAAGAAGTGGGCTCTGGAGACTATGACTCCCAAGGAACCCCTGCTCCGGGATGAAAC 224
Db 116 GAGGAATGGGCTCAGGGAGTATGACTCCATGAGGAACCCGTTTCCGTGAGAAAT 175
QY 225 GTCCATTTCAATAGGATCTTCTCCCAACCATCTACTCATCTCTTTCTGACTGGCAAT 284
Db 176 GCTAATTTCAATAAATCTTCTCCCAACCATCTACTCATCTCTTTCTTAACTGGCAAT 235
QY 285 GTCGCGCATGATTTGCTGCTCCTGCTCATGGTTTACCAGAGAGCTTAAGGAGCATGACG 344
Db 236 GTGGCAATGATTTGCTGCTCCTGCTCATGGTTTACCAGAGAGCTTAAGGAGCATGACG 295
QY 345 GACAAGTACCGGCTGCACCTGTCTGAGTGGCTGACCTCTCTTTGTCTATCACAACCTCCCTTC 404
Db 296 GACAAGTACAGGCTGCACCTGTCTGAGTGGCTGCACCTCTCTTTGTCTATCACAACCTCCCTTC 355

QY 405 TGGGAGTTGATGCCATGGCTGACTGGTACTTTTGGGAAATTTTGTGTAAGGCTGTCCA 464
Db 356 TGGGAGTTGATGCCATGGCTGGCAAACTGGTACTTTGGGAACTTCTATGCAAGGCACTCCA 415
QY 465 ATCATCTACAGTCTCAACCTCTACAGACGGTTCTCATCCTGGCCTTCATCAGCCTGGA 524
Db 416 GTCATCTACAGTCTCAACCTCTACAGACGGTTCTCATCCTGGCCTTCATCAGTCTGGA 475
QY 525 CCGTACTCTGGCCTTCTCCAGCCCAACAGTCAAAAGGCCAAGAACTGCTGGCTGAA 584
Db 476 CCGTACTCTGGCCTTCTCCAGCCCAACAGTCAAAAGGCCAAGAACTGCTGGCTGAA 535
QY 585 AAGGAGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATACCTGACTTCATC 644
Db 536 AAGTGGTCTATGTGGGCGTCTGGATCCAGCCCTCTCTCTGACTATTCCTGACTTCATC 595
QY 645 TTTGGCGAGTCTAGCCAGGGGAGCATCAG 673
Db 596 TTTGGCAAGTCTAGTGAGGAGATGACAG 624

RESULT 3
US-09-023-655-1213
; Sequence 1213, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9219868
US-09-023-655-1213

Query Match 65.1%; Score 445.8; DB 4; Length 1664;
Best Local Similarity 83.0%; Pred. No. 1.9e-116;
Matches 522; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 45 CAGGTGAGGTAGTACAGTACCCCTCTGAGGCGTTTGGTGTCTCGGTAAACACACACGCGTGT 104
Db 1 CGGCAGCAGGTAGTAAAGTACGACCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTGG 60
QY 105 AGAGCAGGTGTTCCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATTAACACTCT 164
Db 61 AGAACACGCGTTACCATGGA-----GGGATCAGTATATACACTTCACTACACTACACC 115
QY 165 GAAGAAAGTGGGCTCTGGAGACTATGATCTCAACAGGAACCCCTGCTTCCGGGATGAAGAAC 224
Db 116 GAGGAAATGGCTCAGGGGACTATGATCTCATGAAGGAACCCCTGTTTCCGTTGAAGAAAT 175
QY 225 GTCCATTCAATAGGATCTTCTGCGCCACCATCTCTCTTGTGATCATCTCTTGTGATGTCATA 284
Db 176 GCTAAATTCATTAATATCTTCTGCGCCACCATCTCTCTTGTGATCATCTCTTGTGATGTCATA 235
QY 285 GTCCGCAATGGATTGGTGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 344
Db 236 GTCCGCAATGGATTGGTGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 295
QY 345 GACAAGTACCGGTGACCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 404
Db 296 GACAAGTACAGGTGACCTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 355
QY 405 TGGGCAATGGATTGGTGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 464
Db 356 TGGGCAATGGATTGGTGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 415
QY 465 ATCATCTACACTGTCAACCTCTACAGAGCGTGTCTCATCTGCGCTTCTATCACTGCGTGGAC 524
Db 416 GTCATCTACAGTCAACCTCTACAGAGCGTGTCTCATCTGCGCTTCTATCACTGCGTGGAC 475
QY 525 CGGTACTCGCCATGTCCAGCGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAA 584
Db 476 CGGTACTCGCCATGTCCAGCGCCACCAAGTCAAGGCGCAAGGAACTGCTGGCTGAA 535
QY 585 AAGCGAGTCTATGTGGCGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 644
Db 536 AAGTGTGTCTATGTGGCGTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 595
QY 645 TTTGCCAGCTCAGCGGCGGAGCATCAG 673
Db 596 TTTGCCAGCTCAGTGAGCGAGATGACAG 624

RESULT 4
US-08-202-056-4
; Sequence 4, Application US/08202056
; Patent No. 5440021
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Hebert, Caroline
; APPLICANT: Kim, Kyung Jin
; APPLICANT: Lee, James
; TITLE OF INVENTION: Antibodies to Human IL-8 Type B Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,056
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/677211
; FILING DATE: 29-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-202-056-4
Query Match 64.1%; Score 439.2; DB 1; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;
QY 50 GCAGGTAGTACGTACCCCTCTGAGCGGTTTGGTGTCTCGGTAAACACACACGCGTCTAGAGC 109
Db 20 CGCGGGCGGCAAGTGAACCGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTCGAGAAC 79
QY 110 GAGTGTGGCCATGAAACCGATCAGTGTGAGTATATACACTTCTGATTAACACTACTCTGAAGA 169
Db 80 CAGCGGTTACCATGGA-----GGGATCAGTATATACACTTCACTACACCGAGGA 134
QY 170 AGTGGGTCTCGAGACTATGATCTCAACAGGAACCCCTGCTTCCGGGATGAAACGTCCTCA 229
Db 135 AATGGGTCTAGGGGACTATGATCTCCATGAAGGAACCCCTGTTTCCGTGAAGAAATGCTAA 194
QY 230 TTTCAATAGGATCTTCTGCGCCACCATCTCTTCTGATCATCTTCTTGTGACTGGCATGTGCG 289
Db 195 TTTCAATAGGATCTTCTGCGCCACCATCTCTTCTGATCATCTTCTTGTGACTGGCATGTGCG 254
QY 230 CAATGGATTTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349
Db 255 CAATGGATTTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
QY 350 GTACGGGTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 409
Db 315 GTACGGGTGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374
QY 410 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469
Db 375 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
Db 435 CTACACTGTCAACCTCTACAGCAGCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
QY 530 CCTGCGCATTTGTCACGCGCCACCAACAGTCAAGGCGCAAGGAACTGCTGCTGCTGCTGCTGCTGCT 589
Db 495 CCTGCGCATTTGTCACGCGCCACCAACAGTCAAGGCGCAAGGAACTGCTGCTGCTGCTGCTGCTGCT 554
QY 590 AGTCTATGTGGCGGTCTGAGTCCAGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
Db 555 GGTCTATGTGGCGGTCTGAGTCCAGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
QY 650 CGACGTGACGCGGCGGACATCAG 673
Db 615 CAACGTGACGCGGCGGACATGACAG 638

RESULT 5
US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:

APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076.093A
FILING DATE: 11-Jun-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-076-093A-3

Query Match 64.1%; Score 439.2; DB 1; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;
QY 50 GCAGGTAGCAGTCAACCTCTGAGGGGTTTGGTGTCCCGGTACCAACACCGGCTGTAGAGC 109
DB 20 GCGGCGCGCAAGTGAAGCGGAGGGGCTGAGTGTCCAGTAGCCACCGCATCTGGAGAAC 79
QY 110 GAGTGTGCCATGGAAACCGATCAGTGTGAGTATATACACTTCTGATCTACTCTGAAGA 169
DB 80 CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACCGAGGA 134
QY 170 AGTGGGGTCTGGAGACTATGACTCCAAAGGAACCTGCTTCCGGGATGAAACGTCCA 229
DB 135 AATGGGCTCAGGGACTATGATCCATGAGGAACCTGTTCCGTGAGGAATGCTAA 194
QY 230 TTTCATAGAGATCTTCCTCCCAACCATCTACTTCTATCTATCTTTGATCTGGCATAGTCGG 289
DB 195 TTTCATAGAGATCTTCCTCCCAACCATCTACTTCTATCTATCTTTAACTGGCATTTGGG 254
QY 290 CAATGATTTGTGATCTCTGCTGATGGTTTACCAAGGAAGCTTAAGAGCATGACGACAA 349
DB 255 CAATGATTTGTGATCTCTGCTGATGGTTTACCAAGGAAGCTTAAGAGCATGACGACAA 314
QY 350 GTACCGGCTGCACCTGTCTAGTGGGTGACCTCTCTTTGTCTATCATCACTCCCTCTGGGC 409
DB 315 GTACAGGCTGCACCTGTCTAGTGGGTGACCTCTCTTTGTCTATCATCACTCCCTCTGGGC 374
QY 410 AGTTGATGCCATGGCTGACTGTGTTGGGAAATTTTGTGTAGGCTGTCCATATCAT 469

DB 375 AGTTGATGCCGTGGCAAACTGGTACTTTGGGAACCTTCCTATGCAAGGCAGTCCATGTGCAT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCGTCTCATCTCTGGCCTTCATCAGCCTGGACCGGTA 529
DB 435 CTACACAGTCAACCTCTACAGCAGTGTCTCATCTCTGGCCTTCATCAGTCTGGACCGCTA 494
QY 530 CCTGGCCATTGTCCACCGCCACCAACAGTCAAGAGGCAAGAACTGCTGGCTGAAAAGGC 589
DB 495 CCTGGCCATCGTCCACCGCCACCAACAGTCAAGAGGCAAGAACTGCTGGCTGAAAAGGT 554
QY 590 AGTCTATGTGGGCGTCTGGATCCAGCGCTCTCTCTGACTATACCTGACTTCTCTTTGC 649
DB 555 GGTCTATGTGGCGTCTGGATCCCTGCCCTCTCTGCTGACTATTCCTCGGACTTCTCTTTGC 614
QY 650 CGAGCTCAGCCAGGGGGACATCAG 673
DB 615 CACGTCAGTGAGGCAGATGACAG 638

RESULT 6
US-08-701-265-3
Sequence 3, Application US/08701265
Patent No. 5776457
GENERAL INFORMATION:
APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,265
FILING DATE: 22-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-Jun-1993
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-701-265-3
Query Match 64.1%; Score 439.2; DB 1; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;

Matches	516;	Conservative	0;	Mismatches	103;	Indels	5;	Gaps	1;
QY	50	GCAGTAGCAGTGACCCCTCTGAGCGGTTGGTGTCCCGTAACCAACCAACCGCTGTGAGAC	109						
Db	20	GCGGCGCAAAAGTGACGCGCGAGGGCGTGTCTCCAGTAGCCACCGCATCTGGAGAAC	79						
QY	110	GAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCGTATATACTCTGTGAGA	169						
Db	80	CAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCAGATAACTACACCGAGA	134						
QY	170	AGTGGGCTCTGGAGACTATGACTCCAAACAGGAACCCCTGCTTCCGGGATGAACCGTCCA	229						
Db	135	AATGGGCTCAGGGACTATGACTCCATGAAGGAACCCCTGTTCCGTGAAGAAATGCTAA	194						
QY	230	TTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCTTCTTGACTGGCGATAGTCGG	289						
Db	195	TTTCAATAAAATCTTCTGCGCCACCATCTACTCCATCATCTTCTTAACGTGCGATTTGTGG	254						
QY	290	CAATGGATTGTGTATCTGTGTCATGGTTTACCAGAAAGAACTAAAGAGCATGAACGACAA	349						
Db	255	CAATGGATTGTCATCTGTCATGGTTTACCAGAAAGAACTAGAGAAGCATGACGACAA	314						
QY	350	GTAACGGGTGCACTGTCAAGTGGTGACCTCTCTTTTGTTCATCACTCCCTCTTCTGGGC	409						
Db	315	GTCAGGGTGACCTGTCAAGTGGCGCACTCTCTTTGTTCATCAGGCTTCCCTTCTGGGC	374						
QY	410	AGTTGATGCCATGCTGACTGGTACTTTTGGGAAATTTTGTGAAGCTGTCCATATCAT	469						
Db	375	AGTTGATGCCGTGGCAACTGTGATCTTTGGGAACCTTCTATGCAAGGCAGTCCATGTCT	434						
QY	470	CTACACTGTCAACCTCTACAGCAGCGTTCTCATCTCTGGCTTTCATCAGCTGGAACCGGTA	529						
Db	435	CTACACAGCTCAACTCTTACAGCAGTGTCTCTCATCTCTGGCTTTCATCAGCTGGACCGCTA	494						
QY	530	CCTGGCATTGTCCACGCCCAACACAGTCAAAAGGCCAAGGAACTGCTGGCTGAAAGGC	589						
Db	495	CCTGGCCATCTGTCCACGCCCAACACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAAGGCT	554						
QY	590	AGTCTATGTGGCGCTCTGGATCCAGCGCCCTCTCTGACTATACCTTGACTTCATCTTTGC	649						
Db	555	GGTCTATGTGGCGTCTGGATCCCTGCGCTCTGCTGACTATCCCGACTTCATCTTTGC	614						
QY	650	CGACTCTACGCGGGGACATCAG	673						
Db	615	CAACGTCAGTGAGGCAGATGACAG	638						

```

RESULT 7
US-08-284-586-3
; Sequence 3, Application US/08284586
; Patent No. 5840856
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinFatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,586
; FILING DATE:

```


LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-802-627A-3

Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

QY 50 GCAGGTAGCAGTACCTCTGAGCGCTTTGGTCTCCGGTACCAACCGGCTGTAGGC 109
Db 20 GCGCGGCGCAAGTAGCAGCGCGGCGCTGTAGTCTCCAGTAGCCACCGCATCTGGAGAC 79

QY 110 GAGTGTTCATCGAAGCCAGTACAGTGTAGTATATACACTTCTGATAACTACTCTGAAGA 169
Db 80 CAGCGGTACCATGGA-----GGGGATCAGTATATACACTTTCAGATAACTACACCGAGA 134

QY 170 AGTGGGCTGTGAGACTATGACTCCAAAGGAACCGCTGCTCCGGGATGAAACGTCCA 229
Db 135 AATGGGCTCAGGAGACTATGACTCCATGAGGAACCGCTGTTCCGTTGAAGAAATGCTAA 194

QY 230 TTTCAATAGATCTTCTGCGCCACCATCTACTTTCATCATCTTCTGACTGGCATAGTGG 289
Db 195 TTTCAATAAATCTTCTGCGCCACCATCTACTTTCATCATCTTCTGACTGGCATAGTGG 254

QY 290 CAATGATTCGATCTCTGCTCATGCTTACCAAGAAAGCTAAGAGCATGACGACAA 349
Db 255 CAATGATTCGATCTCTGCTCATGCTTACCAAGAAAGCTAAGAGCATGACGACAA 314

QY 350 GTACCGGCTGCACCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 409
Db 315 GTACCGGCTGCACCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374

QY 410 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
Db 375 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434

QY 470 CTACACTGTCAACTCTACAGAGCGCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
Db 435 CTACACTGTCAACTCTACAGAGCGCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 494

QY 530 CCTCGCCATTGTCCACGCGCAACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 589
Db 495 CCTCGCCATTGTCCACGCGCAACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 554

QY 590 AGTCTATGCTGGGCTCTGATCCCGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649
Db 615 CAACGTGATGAGGAGATGACAG 638

RESULT 10
US-08-801-238-3
Sequence 3, Application US/08801238
Patent No. 5919896
GENERAL INFORMATION:
APPLICANT: Lee, James
APPLICANT: Wood, William I.
TITLE OF INVENTION: PF4A RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,238
FILING DATE: 19-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-801-238-3

Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

QY 50 GCAGGTAGCAGTACCTCTGAGCGCTTTGGTCTCCGGTAAACCAACCGGCTGTAGGC 109
Db 20 GCGCGGCGCAAGTAGCAGCGCGGCGCTGTAGTCTCCAGTAGCCACCGCATCTGGAGAC 79

QY 110 GAGTGTTCATCGAAGCCAGTACAGTGTAGTATATACACTTCTGATAACTACTCTGAAGA 169
Db 80 CAGCGGTACCATGGA-----GGGGATCAGTATATACACTTTCAGATAACTACACCGAGA 134

QY 170 AGTGGGCTGTGAGACTATGACTCCAAAGGAACCGCTGCTCCGGGATGAAACGTCCA 229
Db 135 AATGGGCTCAGGAGACTATGACTCCATGAGGAACCGCTGTTCCGTTGAAGAAATGCTAA 194

QY 230 TTTCAATAGATCTTCTGCGCCACCATCTACTTTCATCATCTTCTGACTGGCATAGTGG 289
Db 195 TTTCAATAAATCTTCTGCGCCACCATCTACTTTCATCATCTTCTGACTGGCATAGTGG 254

QY 290 CAATGATTCGATCTCTGCTCATGCTTACCAAGAAAGCTAAGAGCATGACGACAA 349
Db 255 CAATGATTCGATCTCTGCTCATGCTTACCAAGAAAGCTAAGAGCATGACGACAA 314

QY 350 GTACCGGCTGCACCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 409
Db 315 GTACCGGCTGCACCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374

QY 410 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
Db 375 AGTTGATGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434

QY 470 CTACACTGTCAACTCTACAGAGCGCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
Db 435 CTACACTGTCAACTCTACAGAGCGCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 494

QY 530 CCTCGCCATTGTCCACGCGCAACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 589
Db 495 CCTCGCCATTGTCCACGCGCAACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 554

QY 590 AGTCTATGCTGGGCTCTGATCCCGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649
Db 615 CAACGTGATGAGGAGATGACAG 638

Db 555 GGTCTATGTCGGCTCTGGATCCCTGCGCTCTGCTGACTATTCGCGACTTCATCTTGC 614
QY 650 CGAGTCAGCCAGGGGACATCAG 673
Db 615 CAAGTCAGTCAGGCGATGACAG 638
RESULT 11
US-08-801-228-3
; Sequence 3, Application US/08801228
; Patent No. 5922541
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
; TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801.228
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2PID3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-801-228-3
Query Match 64.1%; Score 439.2; DB 2; Length 1737;
Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;
QY 50 GCAGGTAGCAGTACCCCTCTGAGCGGTTTGGTGCTCCGGTAACCAACCAACCGGCTGTAGAGC 109
Db 20 GCGCGCGCAAGTGACGCCGAGGCGCTGAGTGCTCCAGTAGCCACCGCATCTGGAGAAC 79
QY 110 GAGTGTTCCTGACGATGATGATATACACTTCTGATTAACACTCTCTGAAGA 169
Db 80 CAGCGGTACCATGGA-----GGGATCAGTATATACACTTCTGATTAACACTCTACCGAGGA 134
QY 170 AGTGGGGTCTGAGAGCTATGACTCCAAACAGGAACCCCTGCTTCGGGATGAAACGTTCCA 229

Db 135 AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCCCTGTTCCGTGAAGAAATGCTAA 194
QY 230 TTTCAATAGGATCTTCCTGCGCCACCATCTACTTTCATCATCTTCTTGAAGTGGCATAGTCGG 289
Db 195 TTTCAATAAATCTTCCTGCGCCACCATCTACTTTCATCATCTTCTTAACTGGCATTTGGG 254
QY 290 CAATGATTTGGTATCTCTGGTTCATGGGTTACAGAGGAAGCTTAAGGAGCATGACGGAACA 349
Db 255 CAATGATTTGGTATCTCTGGTTCATGGGTTACAGAGGAAGCTTAAGGAGCATGACGGAACA 314
QY 350 GTACCGGCTGCACCTGTCAGTGGCTGACCTCTCTTTCATCATCACATCCCTTCTTGGGC 409
Db 315 GTACAGGCTGCACCTGTCAGTGGCCGACCTCTCTTTCATCATCGCTTCCCTTCTGGGC 374
QY 410 AGTTGATGCCATGGCTGACTGGTACTTCTTGGGAAATTTTGTAAAGGTGTCATATCAT 469
Db 375 AGTTGATGCCGTCGCAAACTGGTACTTTTGGGAACTTCTATGCAAGGCGAGTCCATGTCT 434
QY 470 CTACACTGTACACCTCTACAGAGGTTCTCATCTGGCCTTCATCAGCCTGGACCGGTA 529
Db 435 CTACACAGTCAACCTCTACAGAGTGTCTCATCTGGCCTTTCATCAGTCTGGACCGGTA 494
QY 530 CCTCGCCATTGTCCACGCGCCACCAACAGTCAAGGCGCAAGGAACTGCTGGTGAAGGCG 589
Db 495 CCTGGCCATGTCACGCGCCACCAACAGTCAAGGCGCAAGGAACTGCTGGTGAAGGCG 554
QY 590 AGTCTATGTGGCGTCTGGATCCCGACCTCTCTCTGACTATACCTGACTTCTTCTTTC 649
Db 555 GGTCTATGTGGCGTCTGGATCCCGACCTCTCTCTGACTATTCCTGCTGACTTCTTCTTTC 614
QY 650 CGAGCTCAGCCAGGGGACATCAG 673
Db 615 CAACGTCAGTGAGGCGATGACAG 638
RESULT 12
US-09-104-296-3
; Sequence 3, Application US/09104296
; Patent No. 6087475
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104.296
; FILING DATE: 24-June-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/564228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:

NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-104-296-3

Query Match 64.1%; Score 439.2; DB 3; Length 1737;

Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
QY 50 GCAGGTAGCAGTACCCCTCTGAGCGCTTTGGTCTCCGGTAACCAACCGCTGTAGAGC 109
Db 20 GCGCGGCGAAAGTACGCGAGGCGCTGAGTCTCCAGTACCGCAGCATCTGGAGAAC 79
QY 110 GAGTGTGCGATGGAACCGATCAGTGTGAGTATATACATCTTGATACCTCTGAGA 169
Db 80 CAGCGGTACCATGGA-----GGGGATCAGTATATACATCTTCCAGTAACTACACCGAGA 134
QY 170 AGTGGGCTTGGAGACTATGACTCCCAAGAAACCTCTCCGGGATGAAACGTCCA 229
Db 135 AATGGCTCAGGGACTATGACTCCATGAGGACCTCTTCCGTGAGAAATGCTAA 194
QY 230 TTTCATAGATCTTCTGCGCCACCACTACATCTTCTTGAAGTGGATAGTCGG 289
Db 195 TTTCATAAATCTTCTGCGCCACCACTACTCCATCATCTTCTTAACTGGCATTTGGG 254
QY 290 CAATGATGTTGATCCTCTGTCATGTTGTTACCAAGAGAGCTAAGGAGCATACGAGCAA 349
Db 255 CAATGATGTTGATCCTCTGTCATGTTGTTACCAAGAGAGCTAAGGAGCATACGAGCAA 314
QY 350 GTACCGGCTGCACCTGTGAGTGGCTACCTCTCTTGTGATACCTCTTGTGATGAGT 409
Db 315 GTACCGGCTGCACCTGTGAGTGGCTACCTCTCTTGTGATACCTCTTGTGATGAGT 374
QY 410 AGTGTATGCGATGCGTACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 469
Db 375 AGTGTATGCGATGCGTACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 529
Db 435 CTACACTGTCAACCTCTACAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 494
QY 530 CCTCGCATGTCCACGCGCACCAAGTCAAGGCGCAAGAACTGCTGGCTGAAAGGC 589
Db 495 CCTCGCATGTCCACGCGCACCAAGTCAAGGCGCAAGAACTGCTGGCTGAAAGGC 554
QY 590 AGTCTATGTTGGGCTGTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649
Db 555 GGTCTATGTTGGGCTGTGGATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 614
QY 650 CGAGCTGACCGGGGACATCAG 673
Db 615 CAACGTCAGTGAGGCGAGATCAG 638
```

RESULT 13

PCI-US94-06380-2
Sequence 2, Application PC/US9406380
GENERAL INFORMATION:
APPLICANT: Chuntharapai, Anan
APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: K. Jin Kim
APPLICANT: Genentech, Inc.

TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY DISORDERS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06380
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US94-06380-2

Query Match 64.1%; Score 439.2; DB 5; Length 1737;

Best Local Similarity 82.7%; Pred. No. 1.4e-114;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;

```
QY 50 GCAGGTAGCAGTACCCCTCTGAGCGCTTTGGTCTCCGGTAACCAACCGCTGTAGAGC 109
Db 20 GCGCGGCGAAAGTACGCGAGGCGCTGAGTCTCCAGTACCGCAGCATCTGGAGAAC 79
QY 110 GAGTGTGCGATGGAACCGATCAGTGTGAGTATATACATCTTCTGATACCTACTCTGA 169
Db 80 CAGCGGTACCATGGA-----GGGGATCAGTATATACATCTTCCAGTAACTACACCGAGA 134
QY 170 AGTGGGCTCTGGAGACTATGACTCCCAAGAAACCTCTCCGGGATGAAACGTCCA 229
Db 135 AATGGCTCAGGGACTATGACTCCATGAGGAAACCTCTTCCGTGAGAAATGCTAA 194
QY 230 TTTCATAGGATCTTCTGCGCCACCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 289
Db 195 TTTCATAAATCTTCTGCGCCACCACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 254
QY 290 CAATGATGTTGATCCTCTGTCATGTTGTTACCAAGAGAGCTAAGGAGCATACGAGCAA 349
Db 255 CAATGATGTTGATCCTCTGTCATGTTGTTACCAAGAGAGCTAAGGAGCATACGAGCAA 314
QY 350 GTACCGGCTGCACCTGTGAGTGGCTACCTCTCTTGTGATACCTCTCTCTCTCTCTCTCT 409
Db 315 GTACCGGCTGCACCTGTGAGTGGCTACCTCTCTTGTGATACCTCTCTCTCTCTCTCTCT 374
QY 410 AGTGTATGCGATGCGTACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 469
Db 375 AGTGTATGCGATGCGTACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 529
Db 435 CTACACTGTCAACCTCTACAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 494
```

QY 530 CCTCGCCATTGTCACCGCCACCAACAGTCAAGGCGCAAGGAACTGCTGCTGCTGAAAGGC 589
DB 495 CCTGGCCATCGTCCACGCCCAACCAAGTCAAGGCGCAAGGAACTGCTGCTGCTGAAAGGT 554
QY 590 AGTCTATGTTGGCGCTGTTGGATCCCGAGCCCTCCCTGCTGACTATACCTGACTTCACTTTTCG 649
DB 555 GGTCTATGTTGGCGCTGTTGGATCCCGAGCCCTCCCTGCTGACTATTCGCGACTTCACTTTTCG 614
QY 650 CGAGCTCAGCCAGGGGAGCATCAG 673
DB 615 CAAGCTCAGTGAGGCGAGATGACAG 638

RESULT 14

US-09-016-434-1235
; Sequence 1235, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1235:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g189313

US-09-016-434-1235

Query Match 62.9%; Score 430.8; DB 4; Length 1225;
Best Local Similarity 87.5%; Pred. No. 3e-112;
Matches 471; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 136 TGAGTATATACACTTCTGATAACTACTCTGAAGAGTGGGCTCTGGAGACTATGACTCCA 195
DB 36 TCAGTATATACACTTCTGATAACTACTCTGAGGAGAAATGGGCTCAGGGGACTATGACTCCA 95
QY 196 ACAGGACCCCTGCTCCGGGATGAAACGTCCTTTCAATAGATCTTCTCTGCCACCA 255
DB 96 TGAAGGAACCCCTGTTCCGTTGGAAGAAATGCTATTTCAATAAAATCTTCTGCCACCA 155

QY 256 TCTACTTCTCATCATCTTCTTGATCTGGCATATGCGCAATGGATTGGTATCTCTGTCATGG 315
DB 156 TCTACTTCTCATCATCTTCTTAACTGGCATATGTTGGCAATGGATTGGTATCTCTGTCATGG 215
QY 316 GTTACCAAGAAAGTAAAGGAGCATGACGGAACAAGTACCGGCTGCACCTGTCAGTGGCTG 375
DB 216 GTTACCAAGAAAGTAAAGGAGCATGACGGAACAAGTACCGGCTGCACCTGTCAGTGGCTG 275
QY 376 ACCTCTCTTCTTGTCATCATCTCCCTTCTGGGAGTGTGATGCCATGGCTGACTTGGTACT 435
DB 276 ACCTCTCTTCTTGTCATCATCTCCCTTCTGGGAGTGTGATGCCATGGCTGACTTGGTACT 335
QY 436 TTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCG 495
DB 336 TTGGGAACTTCTTATGCAAGGAGTCCATGTCATCTACAGTCAACCTCTACAGCAGTG 395
QY 496 TTCTCATCTCTGGCTTCTATCAGCCTGGAACCGGTACCTGCGCATTTGTCACGCCCAACA 555
DB 396 TCCTCATCTCTGGCTTCTATCAGTCTGGAACCGGTACCTGCGCATTTGTCACGCCCAACA 455
QY 556 GTCAAGGCCCAAGGAAGTCTGCTGGAAGGAGTCTATGTTGGGCTGCTGGATCCCGAG 615
DB 456 GTCAAGGCCCAAGGAAGTCTGCTGGAAGGAGTCTATGTTGGGCTGCTGGATCCCGAG 515
QY 616 CCTCTCTCTGACTATACCTGACTTCTTGTGCGAGCTGAGCCAGGCGGAGCATCAG 673
DB 516 CCTCTCTCTGACTATTCCTGACTTCTTGTGCGAGCTGAGCCAGGCGAGCATGACAG 573

RESULT 15

US-08-153-848-45
; Sequence 45, Application US/08153848
; Patent No. 5759804
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 5759804el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESS: Bicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,848
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5759804and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1317 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; US-08-153-848-45

Query Match 62.9%; Score 430.8; DB 4; Length 1225;
Best Local Similarity 87.5%; Pred. No. 3e-112;
Matches 471; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 136 TGAGTATATACACTTCTGATAACTACTCTGAAGAGTGGGCTCTGGAGACTATGACTCCA 195
DB 36 TCAGTATATACACTTCTGATAACTACTCTGAGGAGAAATGGGCTCAGGGGACTATGACTCCA 95
QY 196 ACAGGACCCCTGCTCCGGGATGAAACGTCCTTTCAATAGATCTTCTCTGCCACCA 255
DB 96 TGAAGGAACCCCTGTTCCGTTGGAAGAAATGCTATTTCAATAAAATCTTCTGCCACCA 155

```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 201...1211
; US-08-153-848-45

Query Match      62.5%; Score 428; DB 1; Length 1317;
Best Local Similarity 87.0%; Pred. No. 1.9e-111; Indels 0; Gaps 0;
Matches 470; Conservative 0; Mismatches 70;

Qy 134 TGTGATATATACACTTCTGATTAACCTCTCTGAAGAGTGGGGTCTGGAGACTATGACTC 193
Db      |||
Qy 164 TTTGCAGATATACACTTCAGATAACTACACCCAGAGAAATGGGCTCAGGGGACTATGACTC 223
Db      |||
Qy 194 CACACAGGACCCCTGCTCCGGGATGAACACGTCCTCAATAGGATCTTCTGCCAC 253
Db      |||
Qy 224 CATGAGGACCCCTGTTCCGTGAAGAAATGCTAAATTTCAATAAATCTTCTGCCAC 283
Db      |||
Qy 254 CATCTACTTCATCATCTTCTTGACTGGCATAGTCCGCAATGGATTGGTGCATCTGGTCA 313
Db      |||
Qy 284 CATCTACTCCATCATCTTCTTAACTGGCATTTGGGCAATGGATTGGTGCATCTGGTCA 343
Db      |||
Qy 314 GGGTTACCAAGAAAGCTAAGGAGCATGACGGACAGTACCCGGCTGCACCTGTCAAGTGC 373
Db      |||
Qy 344 GGGTTACCAAGAAAGCTAAGGAGCATGACGGACAGTACCCGGCTGCACCTGTCAAGTGC 403
Db      |||
Qy 374 TGACCTCCTCTTTGTGATCACAACCTCCCTCTCTGGGAGTTGATGCCATGGCTGACTGTA 433
Db      |||
Qy 404 CGACCTCCTCTTTGTGATCACAACCTCCCTCTCTGGGAGTTGATGCCATGGCTGTA 463
Db      |||
Qy 434 CTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACTCTACAGCAG 493
Db      |||
Qy 464 CTTTGGGAACTTCTATGCAAGGAGTCCATGTCACTACACAGTCAACCTCTACAGCAG 523
Db      |||
Qy 494 CGTTCCTCATCTGGCCCTTCATCAGCTGGACCGGTACCTCGCCATTGTCCAGCCACCAA 553
Db      |||
Qy 524 TGTCCTCATCTGGCCCTTCATCAGCTGGACCGCTACCTGGCCATCGTCCAGCCACCAA 583
Db      |||
Qy 554 CAGTCAAAGGCCAAGGAAACTGTGGCTGAAAGGAGTCTATGTGGGGTCTGGATCCC 613
Db      |||
Qy 584 CAGTCAGAGGCCAAGGAGTGTGGCTGAAAGGAGTCTATGTGGGGTCTGGATCCC 643
Db      |||
Qy 614 AGCCCTCCTCTGACTATACCTGACTTCTATCTTTGGGAGCTCAGCCGGGGACATCAG 673
Db      |||
Qy 644 TGCCCTCCTCTGACTATCTCCGACTTCTATCTTTGCCAACGTCAGTGAGGAGATGACAG 703
Db      |||
```


Db 12 GCGGAGCAGGTAGCAAAAGTAGCGCGGAGGCGCTGAGTGCTCCAGTAGCCACCGCATCTG 71
QY 104 TAGACCGAGTGTGGCCATGGAACCATCAGTGTGAGTATATACACTTCTGATTAATCTACTC 163
Db 72 GAGAACCCAGCGGTTACCATGGA-----GGGATCAGTATATATACACTTCCAGATAACTACAC 126
QY 164 TGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAATGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGCCCCACCATCTACTTCACTTCTTCTGACTGGCAT 283
Db 187 TGCTAAATTTCAATAAAATCTTCTGCCCCACCATCTACTTCCATCATCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 343
Db 247 TGTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 306
QY 344 GGAAGAAGTACCGGCTGCACTGTGAGTGGTGAACCTCTCTTTGTATACACTTCCCTT 403
Db 307 GGAAGAAGTACCGGCTGCACTGTGAGTGGGCAACCTCTCTTTGTATACACTTCCCTT 366
QY 404 CTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 463
Db 367 CTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 426
QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTGCGCTTTCATCAGCCTGGA 523
Db 427 TGTCACTACAGTGTCAACCTCTACAGCAGTGTCTCATCTGCGCTTTCATCAGTCTGGA 486
QY 524 CCGGTACCTCCGATGTGCGGCTGAGTCCAGCCACCAAGTCAAGGCAAGAACTGCTGGCTGA 583
Db 487 CGGTACCTCCGATGTGCGGCTGAGTCCAGCCACCAAGTCAAGGCAAGAACTGCTGGCTGA 546
QY 584 AAAGGAGTCTATGTGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTTCAT 643
Db 547 AAAGTGTCTATGTGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTTCAT 606
QY 644 CTTTCCGAGCTCAGCCAGGGGAGCATCAG 673
Db 607 CTTTCCCAACGTCAGTAGGCGAGATGACAG 636

RESULT 2
US-10-181-906-9
; Sequence 9, Application US/10181906
; Publication No. US20040053864A1
; GENERAL INFORMATION:
; APPLICANT: Karstensy, Gerard
; APPLICANT: Amling, Michael
; APPLICANT: Ducey, Patricia
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA
; TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY
; FILE REFERENCE: 9142-020-999
; CURRENT APPLICATION NUMBER: US/10/181,906
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/02040
; PRIOR FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.0
; SEQ ID NO 9
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-181-906-9

Query Match 65.2%; Score 446.8; DB 13; Length 1679;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGGTGAGGTAGCAGTGACCTCTGAGGCGTGTGGTCTCCGTTAACCACCAAGCGCTG 103
Db 12 GCGGAGCAGGTAGCAAAAGTAGCGCGGAGGCGCTGAGTGCTCCAGTAGCCACCGCATCTG 71

QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAAATCTACTC 163
Db 72 GGAACCCAGCGGTTACCATGGA-----GGGATCAGTATATATACACTTCCAGATAACTACAC 126
QY 164 TGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGCCCCACCATCTACTTCACTTCTTCTGACTGGCAT 283
Db 187 TGCTAAATTTCAATAAAATCTTCTGCCCCACCATCTACTTCCATCATCTTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 343
Db 247 TGTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 306
QY 344 GGAAGAAGTACCGGCTGCACTGTGAGTGGTGAACCTCTCTTTGTATACACTTCCCTT 403
Db 307 GGAAGAAGTACCGGCTGCACTGTGAGTGGGCAACCTCTCTTTGTATACACTTCCCTT 366
QY 404 CTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 463
Db 367 CTGGGCAATGGATTGGTGATCTCTGGTCTATGGGTTTACCAGAAAGCTTAAGGAGCATGAC 426
QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTGCGCTTTCATCAGCCTGGA 523
Db 427 TGTCACTACAGTGTCAACCTCTACAGCAGTGTCTCATCTGCGCTTTCATCAGTCTGGA 486
QY 524 CCGGTACCTCCGATGTGCGGCTGAGTCCAGCCACCAAGTCAAGGCAAGAACTGCTGGCTGA 583
Db 487 CGGTACCTCCGATGTGCGGCTGAGTCCAGCCACCAAGTCAAGGCAAGAACTGCTGGCTGA 546
QY 584 AAAGGAGTCTATGTGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTTCAT 643
Db 547 AAAGTGTCTATGTGGGCTGTGATCCAGCCCTCTCTGACTATACCTGACTTTCAT 606
QY 644 CTTTCCGAGCTCAGCCAGGGGAGCATCAG 673
Db 607 CTTTCCCAACGTCAGTAGGCGAGATGACAG 636

RESULT 3
US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-912
Query Match 65.2%; Score 446.8; DB 13; Length 1679;

Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGGTGCAAGTACAGTACCTCTGAGGCGTTTGGTCTCCGGTAACCAACCGGCTG 103
Db 12 GCGGAGCAGGTAGCAAGTACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGTATACTCTC 163
Db 72 GAGAACCGCGGTTACCATGGA-----GGGATCAGTATATACACTTTCAGATAACTACAC 126
QY 164 TGAAGAAGTGGGCTCGGAGACTATGACTCCAAACAGGAACCGTCTCCGGGATGAAA 223
Db 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCGTGTTCGGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGCCACCACTACTTCTCATCATCTTCTGACTGGCAT 283
Db 187 TGTAAATTTCAATAAATCTTCTGCCACCATCTACTTCTTAACTGGCAT 246
QY 284 AGTCGCAATGGAATGGTGTATCTGTGTCATGAGTTACCAAGAAAGCTAAGGAGCATGAC 343
Db 247 TGTGGCAATGGATTTGTCTATCTGTGTCATGGGTTACCAAGAAACTGAGAAGCATGAC 306
QY 344 GGCAGATACCGCTCACCTGTCAGTGGTGCACCTCTTGTGTCATCACATCCCTTT 403
Db 307 GGCAGATACAGGCTGCACCTGTCAGTGGGCGACCTCTTGTGTCATCACGCTTCCCTT 366
QY 404 CTGGGAGTTGATGCCATGCTGACTGCTGACTTTGGGAAATTTTGTGAAGCTGTCCA 463
Db 367 CTGGGAGTTGATGCCATGCTGCAACTGGTACTTTGGGAACTTCTTATGCAAGGAGTCCA 426
QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCTTCTATCTGCTGGCTTCTATCAGCTGGA 523
Db 427 TGTATCTACAGTCAACCTCTACAGCAGTGTCTCTTCTGCTGGCTTCTATCAGTCTGGA 486
QY 524 CCGGTACTCGCATTTGTCAGCCCAACAGTCAAGGCGCAAGGAACTGCTGGCTGA 583
Db 487 CCGGTACTCGCATTTGTCAGCCCAACAGTCAAGGCGCAAGGAACTGCTGGCTGA 546
QY 584 AAGGAGTCTATGTGGGCTGTGATGCCAGCCCTCTCTGACTATATCTGACTTTCAT 643
Db 547 AAGGAGTCTATGTGGGCTGTGATGCCAGCCCTCTCTGACTATATCTGACTTTCAT 606
QY 644 CTTTGGCGAGTCTGAGCCAGGGGAGCATCAG 673
Db 607 CTTTGGCAACGTCTAGTGAGGAGCATGACAG 636

RESULT 4

US-10-151-274-14
; Sequence 14, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-151-274-14

Query Match 65.2%; Score 446.8; DB 13; Length 1679;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGGTGCAAGTACAGTACCTCTGAGGCGTTTGGTCTCCGGTAACCAACCGGCTG 103
Db 12 GCGGAGCAGGTAGCAAGTACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGTATACTCTC 163
Db 72 GAGAACCGCGGTTACCATGGA-----GGGATCAGTATATACACTTTCAGATAACTACAC 126
QY 164 TGAAGAAGTGGGCTCGGAGACTATGACTCCAAACAGGAACCGTCTCCGGGATGAAA 223
Db 127 CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGGAACCGTGTTCGGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGCCACCACTACTTCTCATCATCTTCTTGAAGTGGCAT 283
Db 187 TGTAAATTTCAATAAATCTTCTGCCACCATCTACTTCTTAACTGGCAT 246
QY 284 AGTCGCAATGGAATGGTGTATCTGTGTCATGAGTTACCAAGAAAGCTAAGGAGCATGAC 343
Db 247 TGTGGCAATGGATTTGTCTATCTGTGTCATGGGTTACCAAGAAACTGAGAAGCATGAC 306
QY 344 GGCAGATACCGGCTCACCTGTCAGTGGTGCACCTCTTGTGTCATCACATCCCTTT 403
Db 307 GGCAGATACAGGCTGCACCTGTCAGTGGGCGACCTCTTGTGTCATCACGCTTCCCTT 366
QY 404 CTGGGAGTTGATGCCATGCTGACTGCTGACTTTGGGAAATTTTGTGAAGCTGTCCA 463
Db 367 CTGGGAGTTGATGCCATGCTGCAACTGGTACTTTGGGAACTTCTTATGCAAGGAGTCCA 426
QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCTTCTATCTGCTGGCTTCTATCAGCTGGA 523
Db 427 TGTATCTACAGTCAACCTCTACAGCAGTGTCTCTTCTGCTGGCTTCTATCAGTCTGGA 486
QY 524 CCGGTACTCGCATTTGTCAGCCCAACAGTCAAGGCGCAAGGAACTGCTGGCTGA 583
Db 487 CCGGTACTCGCATTTGTCAGCCCAACAGTCAAGGCGCAAGGAACTGCTGGCTGA 546
QY 584 AAGGAGTCTATGTGGGCTGTGATGCCAGCCCTCTCTGACTATATCTGACTTTCAT 643
Db 547 AAGGAGTCTATGTGGGCTGTGATGCCAGCCCTCTCTGACTATATCTGACTTTCAT 606
QY 644 CTTTGGCGAGTCTGAGCCAGGGGAGCATCAG 673
Db 607 CTTTGGCAACGTCTAGTGAGGAGCATGACAG 636

RESULT 5

US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens


```

; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-75

Query Match      65.2%; Score 446.8; DB 15; Length 1679;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGTCAGTACAGTACCTCTGAGCGGTTGGTGTCCGGTAACACACCGCTG 103
Db 12 GCGGCAGCAGGTAGCAAGTACGCGCGGCTGAGTGTCTCCAGTAGCCACCGCATCG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTAGTATATACACTTCTGATACTCTC 163
Db 72 GAGAACACGCGGTACATGGA-----GGGATCAGTATATACACTTTCAGATACTAC 126
QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCAAAGAACCCCTGCTCCGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGTCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 283
Db 187 TGCTAAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGTCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 283
Db 187 TGCTAAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 247 TGTGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 247 TGTGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
QY 344 GGCACAGTACCGGCTGACCTGTGTCAGTGGTGAACCTCTCTTTGTCATCATCTCCCTT 403
Db 307 GGCACAGTACCGGCTGACCTGTGTCAGTGGTGAACCTCTCTTTGTCATCATCTCCCTT 366
QY 404 CTGGGCAAGTGTGATGCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 367 CTGGGCAAGTGTGATGCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTCTGGCCCTTTCATCAGCTTGA 523
Db 427 TGTCTATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTCTGGCCCTTTCATCAGCTTGA 486
QY 524 CCGGTACCTCGCATTTGTCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGGCTGA 583
Db 487 CCGGTACCTCGCATTTGTCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGGCTGA 546
QY 584 AAAGGCAAGTGTATGTGGGCTGCTGATCCCAAGCCCTCTCTCTGACTATACCTGACTTTCAT 643
Db 547 AAAGGCAAGTGTATGTGGGCTGCTGATCCCAAGCCCTCTCTCTGACTATACCTGACTTTCAT 606
QY 644 CTTTGGCCAGTGTACCCAGGGGACATCAG 673
Db 607 CTTTGGCCAGTGTACCCAGGGGACATCAG 636

RESULT 8
US-10-021-660-58
; Sequence 58, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; FILE REFERENCE: 9U 204 205 RI
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18

```

```

; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-58

Query Match      65.2%; Score 446.8; DB 15; Length 1679;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGTCAGTACAGTACCTCTGAGCGGTTGGTGTCCGGTAACACACCGCTG 103
Db 12 GCGGCAGCAGGTAGCAAGTACGCGCGGCTGAGTGTCTCCAGTAGCCACCGCATCTG 71
QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTAGTATATACACTTCTGATACTCTC 163
Db 72 GAGAACACGCGGTACATGGA-----GGGATCAGTATATACACTTTCAGATACTAC 126
QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCAAAGAACCCCTGCTCCGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGTCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 283
Db 187 TGCTAAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 224 CGTCCATTTCAATAGGATCTTCTGTCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 283
Db 187 TGCTAAATTTCAATAAATCTTCTGCGCCACCATCTACTCCATCATCTCTTAACTGGCAT 246
QY 284 AGTCGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 127 CGAGGAAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAAGAAA 186
QY 247 TGTGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 247 TGTGGCAATGGATTTGGTGTGATCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
QY 344 GGCACAGTACCGGCTGACCTGTGTCAGTGGTGAACCTCTCTTTGTCATCATCTCCCTT 403
Db 307 GGCACAGTACCGGCTGACCTGTGTCAGTGGTGAACCTCTCTTTGTCATCATCTCCCTT 366
QY 404 CTGGGCAAGTGTGATGCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
Db 367 CTGGGCAAGTGTGATGCTGCTGATGAGTATGACTCTCAACAGGAACCCCTGCTTCCGGGATGAAAA 223
QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTCTGGCCCTTTCATCAGCTTGA 523
Db 427 TGTCTATCTACAGTGTCAACCTCTACAGCAGGCTTCTCATCTCTGGCCCTTTCATCAGCTTGA 486
QY 524 CCGGTACCTCGCATTTGTCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGGCTGA 583
Db 487 CCGGTACCTCGCATTTGTCCACGCCACCAACAGTCAAGGCCAAGAACTGCTGGCTGA 546
QY 584 AAAGGCAAGTGTATGTGGGCTGCTGATCCCAAGCCCTCTCTCTGACTATACCTGACTTTCAT 643
Db 547 AAAGGCAAGTGTATGTGGGCTGCTGATCCCAAGCCCTCTCTCTGACTATACCTGACTTTCAT 606
QY 644 CTTTGGCCAGTGTACCCAGGGGACATCAG 673
Db 607 CTTTGGCCAGTGTACCCAGGGGACATCAG 636

RESULT 9
US-10-341-434-177
; Sequence 177, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 RI
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18

```

;; PRIORITY APPLICATION NUMBER: US 60/348,164
;; PRIORITY FILING DATE: 2002-01-15
;; PRIORITY APPLICATION NUMBER: US 60/348,119
;; PRIORITY FILING DATE: 2002-01-15
;; NUMBER OF SEQ ID NOS: 238
;; SOFTWARE: Patent in version 3.1
;; SEQ ID NO 177
;; LENGTH: 1679
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (89)..(1144)
;; OTHER INFORMATION:
US-10-341-434-177

Query Match 65.2%; Score 446.8; DB 16; Length 1679;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY	44	GCAGGTGACGAGTACGACCCCTCTGAGGCGTTGGTGTCCCGGTAAACCAACCGGCTG	103
DB	12	CGCGCAGCAGGTAGCAAAAGTACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG	71
QY	104	TAGAGCGAGTGTGGCATGGACCGCATCAGTGTGAGTATATACACTTCTGTAATCTACTC	163
DB	72	GAGAACCCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCCAGATAACTACAC	126
QY	164	TGAAGAATGGGGTCTGGAGACTATGACTCCAAAGAAACCCCTGCTTCCGGGATGAAAA	223
DB	127	CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCCCTGTTCCGTGAAGAAAA	186
QY	224	CGTCCATTTCATAGGATCTTCTGCCCAACCATCTACTTCTATCTCTTCTGATGCGAT	283
DB	187	TGCTAAATTCATAAATCTTCTGCCCAACCATCTACTTCTATCTCTTCTTAACTGGCAT	246
QY	284	AGTCGGCAATGATGGGTGATCTCTGGTCTATGGTGTACCAAGAAAGCTTAAGGACATGAC	343
DB	247	TGTGGCAATGATGGGTGATCTCTGGTGTACCAAGAAAGCTTAAGGACATGAC	306
QY	344	GGACAAGTACCGGTGACACCTCTGAGTGGGTGACCTCTCTTTGTCATCACTCCCTTT	403
DB	307	GGACAAGTACAGGTGACACCTCTGAGTGGGTGACCTCTCTTTGTCATCACTCCCTTT	366
QY	404	CTGGGCGATTGATGCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGCCA	463
DB	367	CTGGGCGATTGATGCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGCCA	426
QY	464	TATCATCTACAGTGTCAACCTCTACAGCAGGTTCTCATCTCTGGCTTCTACGCTTGA	523
DB	427	TGTCACTACAGTGTCAACCTCTACAGCAGGTTCTCATCTCTGGCTTCTACGCTTGA	486
QY	524	CGGTACCTCGCATTTGTCAGCCCAACCAAGTCAAGGCGCAAGGAACTGCTGGCTGA	583
DB	487	CGGTACCTCGCATTTGTCAGCCCAACCAAGTCAAGGCGCAAGGAACTGCTGGCTGA	546
QY	584	AAAGGCGATCTATGTGGGCGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTTCAT	643
DB	547	AAAGTGTCTATGTGGGCGTCTGGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	606
QY	644	CTTTGCCAGCTGACGCGAGGGGACATCAG	673
DB	607	CTTTGCCAGCTGACGCGAGGGGACATCAG	636

RESULT 10
US-09-971-392-20
; Sequence 20, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.

;; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
;; FILE REFERENCE: PA-0029 US
;; CURRENT APPLICATION NUMBER: US/09/971,392
;; CURRENT FILING DATE: 2001-10-03
;; PRIORITY APPLICATION NUMBER: 60/237,652
;; PRIORITY FILING DATE: 2000-10-03
;; NUMBER OF SEQ ID NOS: 260
;; SOFTWARE: PERL Program
;; SEQ ID NO 20
;; LENGTH: 1711
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; OTHER INFORMATION: Template ID: 284616.2
;; NAME/KEY: unsure
;; LOCATION: 1706
;; OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-20

Query Match 65.2%; Score 446.8; DB 10; Length 1711;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY	44	GCAGGTGACGAGTACGACCCCTCTGAGGCGTTGGTGTCCCGGTAAACCAACCGGCTG	103
DB	47	CGCGCAGCAGGTAGCAAAAGTACGCGGAGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTG	106
QY	104	TAGAGCGAGTGTGGCATGGACCGCATCAGTGTGAGTATATACACTTCTGTAATCTACTC	163
DB	107	GAGAACCCAGCGGTTACCATGGA-----GGGGATCAGTATATACACTTCCAGATAACTACAC	161
QY	164	TGAAGAATGGGGTCTGGAGACTATGACTCCAAAGAAACCCCTGCTTCCGGGATGAAAA	223
DB	162	CGAGGAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCCCTGTTCCGTGAAGAAAA	221
QY	224	CGTCCATTTCATAGGATCTTCTGCCCAACCATCTACTTCTATCTCTTCTGACTGGCAT	283
DB	222	TGCTAAATTCATAAATCTTCTGCCCAACCATCTACTTCTATCTCTTCTTAACTGGCAT	281
QY	284	AGTCGGCAATGATGGGTGATCTCTGGTGTACCAAGAAAGCTTAAGGACATGAC	343
DB	282	TGTGGCAATGATGGGTGATCTCTGGTGTACCAAGAAAGCTTAAGGACATGAC	341
QY	344	GGACAAGTACCGGTGACACCTCTGAGTGGGTGACCTCTCTTTGTCATCACTCCCTTT	403
DB	342	GGACAAGTACAGGTGACACCTCTGAGTGGGTGACCTCTCTTTGTCATCACTCCCTTT	401
QY	404	CTGGGCGATTGATGCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCA	463
DB	402	CTGGGCGATTGATGCCATGGCTGACTGTGTTGGGAAATTTTGTGTAAGGCTGTCA	461
QY	464	TATCATCTACAGTGTCAACCTCTACAGCAGGTTCTCATCTCTGGCTTCTACGCTTGA	523
DB	462	TGTCACTACAGTGTCAACCTCTACAGCAGGTTCTCATCTCTGGCTTCTACGCTTGA	521
QY	524	CGGTACCTCGCATTTGTCAGCCCAACCAAGTCAAGGCGCAAGGAACTGCTGGCTGA	583
DB	522	CGGTACCTCGCATTTGTCAGCCCAACCAAGTCAAGGCGCAAGGAACTGCTGGCTGA	581
QY	584	AAAGGCGATCTATGTGGGCGTCTGGATCCAGCCCTCTCTGACTATACCTGACTTTCAT	643
DB	582	AAAGTGTCTATGTGGGCGTCTGGATCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	641
QY	644	CTTTGCCAGCTGACGCGAGGGGACATCAG	673
DB	642	CTTTGCCAGCTGACGCGAGGGGACATCAG	671

RESULT 11
US-10-101-510-459
; Sequence 459, Application US/10101510
; Publication No. US20030148295A1

```

; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 459
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1706)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-459

```

```

Query Match 65.2%; Score 446.8; DB 15; Length 1711;
Best Local Similarity 83.0%; Pred. No. 2.5e-132;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 44 GCAGGTGCAGGTAGCAGTACCTCTGAGCGGTTTGGTGTCCGGTAACACACCGGCTG 103
DB 47 GCGGACGAGGTAGCAAGTGAAGCGGCGCTGAGTGTCCAGTAGCCACCGCATCTG 106

QY 104 TAGAGCGAGTGTGCCATGGAACCGATCAGTGTGAGTATATACACTTGTATATCTC 163
DB 107 GAGAACGAGCGGTATCCATGGA-----GGGGATCAGTATATATACACTTCAATATAC 161

QY 164 TGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAGAGAACCTGTCTCCGGGATGAAA 223
DB 162 CGAGGAATGGGCTCAGGAGCTATGACTCCATGAAGAACCTGTCTCCGTGAGAAA 221

QY 224 CGTCCATTTCAATAGGATCTTCTGCCACCACTTCTCATCTTCTTCTGACTGGCAT 283
DB 222 TGCTAATTTCAATAAATCTTCTGCCACCACTTCTCATCTTCTTCTTAACTGGCAT 281

QY 284 AGTCGGCAATGGATGGTGTATCTGTCTGATGCTGATGCTGATGCTGATGCTGATG 343
DB 282 TGTTGGCAATGGATGGTGTATCTGTCTGATGCTGATGCTGATGCTGATGCTGATG 341

QY 344 GGAACAAGTACCGGCTGCACCTGTGAGTGGTGTGACCTCTCTTTGTCTATCAGCTCC 403
DB 342 GGAACAAGTACAGGCTGCACCTGTGAGTGGTGTGACCTCTCTTTGTCTATCAGCTCC 401

QY 404 CTGGCAGTTGATGCGGCTGGAACCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCA 463
DB 402 CTGGCAGTTGATGCGGCTGGAACCTGGTACTTTGGGAAATTTTGTGTAAGGCTGTCA 461

QY 464 TATCATCTACAGTGTCAACCTCTACAGCAGCGTTCTCATCTGCGCTTCTCATGAGCTGA 523
DB 462 TGTCATCTACAGTGTCAACCTCTACAGCAGCGTTCTCATCTGCGCTTCTCATGAGCTGA 521

QY 524 CCGGTACTCTGCCATTTGTCACGCCACCAACAGTCAAGGCCCAAGGAAATCTGTGGCTGA 583
DB 522 CCGGTACTCTGCCATTTGTCACGCCACCAACAGTCAAGGCCCAAGGAAATCTGTGGCTGA 581

QY 584 AAAGGAGTCTATGTGGGCTGTGGATCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
DB 582 AAAGGAGTCTATGTGGGCTGTGGATCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 641

QY 644 CTTTGGCGAGCTCAGCCAGGCGGAGCATCAG 673
DB 642 CTTTGGCGAGCTCAGTGAGGCGAGATGACAG 671

```

```

RESULT 12
US-10-641-643-1213
; Sequence 1213, Application US/10641643

```

```

; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9219868
; SEQUENCE DESCRIPTION: SEQ ID NO: 1213 :
US-10-641-643-1213

```

```

Query Match 65.1%; Score 445.8; DB 17; Length 1664;
Best Local Similarity 83.0%; Pred. No. 5.2e-132;
Matches 522; Conservative 0; Mismatches 102; Indels 5; Gaps 1;

QY 45 CAGTGTGAGGTAGCAGTACCTCTGAGCGGTTTGGTGTCCGGTAACACACCGGCTGT 104
DB 1 CCGCAGCAGGTAGCAAGTGAAGCGGCGCTGAGTGTCTCCAGTAGCCACCGCATCTGG 60

QY 105 AGACGGAGTGTGCCATGGACCGCATCAGTGTGAGTATATACACTTCTGATACTCTCT 164
DB 61 AGAACCGCGGTTTACCATGGA-----GGGGATCAGTATATATACACTTCCAGATACTAC 115

QY 165 GAAGAAGTGGGCTCTGGAGACTATGACTCCAAAGGAACCCCTCTTCCGGGATGAAAA 224
DB 116 GAGGAATGGCTCAGGGGACTATGACTCCCATGAAGGAACCCCTCTTCCGGTGAAGAAAT 175

QY 225 GTCCATTTCAATAGATCTTCTGCGCCACCATCTACTTCTCATCTCTTCTGACTGGCATA 284
DB 176 GCTAATTTCAATAAATCTTCTGCGCCACCATCTACTTCTCATCTCTTCTTAACTGGCAT 235

QY 285 GTCGCAATGATTGGTGTGATCTCTGTCATGGTTTACCAGAAAGAGTCAAGGAGCATGACG 344
DB 236 GTGGCAGTGGTGTGATCTCTGTCATGGTTTACCAGAAAGTCAAGGAGCATGACG 295

QY 345 GACAAGTACCGGCTGACCTGTCTGAGTGGTGAACCTCTCTTTTGTGTCATCAGTCCCTTC 404
DB 296 GACAAGTACAGGCTGCACCTGTCTGAGTGGCGACCTCTCTTTTGTGTCATCAGCTTCCCTTC 355

```

	Query Match	64.1%	Score 439.2	DB 9	Length 1737	
	Best Local Similarity	82.7%	Pred. No. 7e-130			
	Matches 516	Conservative	0	Mismatches 103	Indels 5	Gaps 1
Qy	50	GCAGGTAGCAGTCACCGCTCTGAGGGCGTTGGTGCTCCGGTAAACACACACCGGCTCTAGAGC	109			
Db	20	GCGCGCCGAAATGATGACGCCGAGGGCCCTGAGTGTCTCCAGTAGCCACCGCATCTGAGAAAC	79			
Qy	110	GAGTGTTCGCATGGAAACCGCATCAGTGTGAGTATATACACTTCTGATACTACTCTGAGA	169			
Db	80	CAGCGGTACATGGA-----GGGGATCAGTATATATACACTTCAGTAACTACACCGAGGA	134			
Qy	170	AGTGGGCTCTGGAGCATATGATCTCCACAAAGAAACCTCTCCGGGATGAAGACGTCCA	229			
Db	135	AATGGGCTCAGGGGACATGATCTCCATGAAAGAACCTTCTTCGTGAAGAAATGTCTAA	194			
Qy	230	TTTCAATPAGGATCTTCTGCGCCACCATCTACTTTCATCTCTTTTGTATGGCATAGTCGG	289			

Qy	290 CAATGGATTGGTGATCCTGTCATGGGTACAGAGAAGCTAAGGAGCATGACGGACAA	319
Db	255 CAATGGATTGGTGATCCTGTCATGGGTACAGAGAAGCAACTGAGAAGCATGACGGACAA	314

QY	350	GTACGGGCTGCACTGTTCAGTGCTGACCTCTCTTTGTTCATCACACTCCCTCTTCGGC	409
Db	315	GTACAGGCTGCACTGTTCAGTGCGACCTCTCTTTGTTCATCAGCTCCCTCTTCGGC	374
QY	410	AGTTGATGCCATGCTGACTGGTACTTTGGGAAATTTTGTGAAGGTGTCCTATATCAT	469
Db	375	AGTTGATGCCGTGGCAACTGTGTACTTTGGAACTTCCTATGCAAGGCACTCCATGTCA	434
QY	470	CTACACGTGTAACTCTTACAGCAGCGTTCTCATCTGGCGCTTCATCAGCGTGGACCGGTA	529
Db	435	CTACACAGTCAACTCTTACAGCAGTGTCTCATCTCGGCTTTTCATCAGTCTGGACCGGTA	494

584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

```

DB      615  CAACGTCAGTGGGACATGACAG 638

RESULT 14
US-10-666-689-3
; Sequence 3, Application US/10666689
; Publication No. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Jee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P0706P2C2D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19

```

RESULT 14

US-10-666-689-3

; Sequence 3, Application US/10666689

Publication No. US20040037830A1
GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Lee, James

APPLICANT: Wood, William I.

TITLE OF INVENTION: Human PF4A Receptors, Nucleic Acid Encoding and

; TITLE OF INVENTION: Antibodies Binding Thereto
 ; FILE REFERENCE: P0706P2C2D2C1

FILE REFERENCE: F0700F02Z01
CURRENT APPLICATION NUMBER: US/10/666,689

;
CURRENT FILING DATE: 2003-09-19

; PRIOR APPLICATION NUMBER: US 09/104,063

;; PRIOR FILING DATE: 1988-06-24
;; PRIOR APPLICATION NUMBER: US 08/701,265

PRIOR FILING DATE: 1996-08-22

; PRIOR APPLICATION NUMBER: US 08/664,228

; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: US 08/078,093
 ; PRIOR FILING DATE: 1993-06-11

PRIOR APPLICATION NUMBER: US 07/810,782

; PRIOR FILING DATE: 1991-12-19

;; PRIOR APPLICATION NUMBER: US 07/677,211
;; PRIOR FILING DATE: 1991-03-29
;; NUMBER OF SEQ ID NOS: 6
;; SEQ ID NO 3
;; LENGTH: 1737
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-10-666-689-3

Query Match 64.1%; Score 439.2; DB 13; Length 1737;
Best Local Similarity 82.7%; Pred. No. 7e-130;
Matches 516; Conservative 0; Mismatches 103; Indels 5; Gaps 1;
QY 50 GCAGGTAGCAGTGCACCTCTGAGCGCTTGGTCTCCGGTAACACACCGCTGTAGAGC 109
DB 20 GCACGGGCGAAAGTGCACCGCGAGGCTGAGTCTCCAGTAGCCACCGCATCTGGAGAAC 79
QY 110 GAGTGTTCATGGAACCGCATGAGTGTAGTATATACACTTCTGATTAATCTCTGAAGA 169
DB 80 CAGCGGTTACCATGGA-----GGGATCAGTATATACACTTTCAGATAATACACCGGAGGA 134
QY 170 AGTGGGCTCGGAGACTATGACTCTCCACAGGAGCCCTGCTCCGGGATGAAACCTCCA 229
DB 135 AATGGGCTCAGGGGACTATGACTCCATGAAGGAACCTGTTCCGGTGAAGAAATCTAA 194
QY 230 TTTCAATAGGATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTGACTGGCATAGTCGG 289
DB 195 TTTCAATAAATCTTCTGCCCCACCATCTACTTTCATCATCTTCTTAAGTGGCAATGTGGG 254
QY 290 CAATGGATTGGTATCTCTGTGTCATGAGTACCGAGAGAGCTAAGAGCATGACGACAA 349
DB 255 CAATGGATTGGTATCTCTGTGTCATGAGTACCGAGAGAGCTAAGAGCATGACGACAA 314
QY 350 GTACCGGCTGCACCTGTGAGTGGCTGACCTCTTGTTCATCACACTCCCTTCTGGGC 409
DB 315 GTACAGGCTGCACCTGTGAGTGGCTGACCTCTTGTTCATCACACTCCCTTCTGGGC 374
QY 410 AGTTGATGCCATGGCTGACTGGTACTTGTGGAAATTTTGTGTAAGCTGTCCATATCAT 469
DB 375 AGTTGATGCCGTGGCAACTGGTACTTGTGGAAATTTTGTGTAAGCTGTCCATATCAT 434
QY 470 CTACACTGTCAACCTCTACAGCAGCTTCTCATCTGGCTTCTCATCAGCTGACCGGTA 529
DB 435 CTACAGCTCAACCTCTACAGCAGCTTCTCATCTGGCTTCTCATCAGCTGACCGGTA 494
QY 530 CCTCGCATTTGTCCAGCCCAACAGTCAAGGCGCAAGAAATCTGCTGGCTGAAAGGC 589
DB 495 CCTGGCCATCTGTCCAGCCCAACAGTCAAGGCGCAAGAAATCTGCTGGCTGAAAGGT 554
QY 590 AGTCTATGTGGGCTGTGGATCCAGGCTTCTCTGACTATACCTGACTTCTTCTTTC 649
DB 555 GGTCTATGTGGGCTGTGGATCCAGGCTTCTCTGACTATACCTGACTTCTTCTTTC 614
QY 650 CGAGCTCAGCCAGGGGACATCAG 673
DB 615 CAAGCTCAGTGAGGCAGATGACAG 638

RESULT 15
US-09-880-107-2143
;; Sequence 2143, Application US/09880107
;; Patent No. US20020142981A1
;; GENERAL INFORMATION:
;; APPLICANT: Horne, Darci T.
;; APPLICANT: Vockley, Joseph G.
;; APPLICANT: Scherf, Uwe
;; APPLICANT: Gene Logic, Inc.
;; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;; FILE REFERENCE: 44921-5028-WO
;; CURRENT APPLICATION NUMBER: US/09/880,107
;; CURRENT FILING DATE: 2001-06-14
;; PRIOR APPLICATION NUMBER: US 60/211,379
;; PRIOR FILING DATE: 2000-06-14

;; PRIOR APPLICATION NUMBER: US 60/237,054
;; PRIOR FILING DATE: 2000-10-02
;; NUMBER OF SEQ ID NOS: 3950
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 2143
;; LENGTH: 1670
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L06797
US-09-880-107-2143

Query Match 63.9%; Score 437.6; DB 9; Length 1670;
Best Local Similarity 83.7%; Pred. No. 2.2e-129;
Matches 509; Conservative 0; Mismatches 94; Indels 5; Gaps 1;
QY 56 CTCGAGCGCTTGGTCTCCGGTAAACACACCGCTGTAGAGCGAGTGTTCCTCATGAA 125
DB 18 CCGCAGGCGCTGAGTGTCTCAGTAGCCACCGCATCTGGAGAACAGCGGTTCACATGGA 77
QY 126 CGGATCAGTGTGATATATACACTTCTGATAAATCTCTGAAGAAAGTGGGTCTGGAGAC 185
DB 78 -----GGGATCAGTATATACACTTTCAGATAAATCTACCGAGGAAATGGGCTCAGGAGAC 132
QY 186 TATGATCCAAACAGGAACCCCTGCTTCGGGATGAAACGTCCTATTCATATAGATCTTC 245
DB 133 TATGATCCATGAAGGAACCCCTGTTCCGTGAAGAAATGCTAATTTCAATAAATCTTC 192
QY 246 CTGCCCCACCTACTTCTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTATC 305
DB 193 CTGCCCCACCTACTTCTCATCATCTTCTTAAGTGGCATTTGGGCAATGGATTGGTATC 252
QY 306 CTGTCATGAGTGTACCAAGAAAGCTAAGAGCATGACGCAAGATACCGGCTGACCTG 365
DB 253 CTGTCATGAGTGTACCAAGAAAGCTAAGAGCATGACGCAAGATACCGGCTGACCTG 312
QY 366 CAGTGGCTGACCTCTCTTGTCTATCATCATCTCCCTTCTGGGAGTTGATGCTATGGCT 425
DB 313 TCAGTGGCGGACCTCTCTTGTCTATCATCAGCTTCCCTTCTGGGAGTTGATGCTATGGCA 372
QY 426 GACTGCTACTTTGGGAAATTTTGTGTAAGGCTGTCCATATCATCTACATGTCAACCTC 485
DB 373 AACTGGTACTTTGGGAACTTCTTATGCAAGGAGTCCATGTCTATCACAGTCAACCTC 432
QY 486 TACAGCAGCTTCTCATCTCGGCTTCTATCAGCTGAGCCGCTACCTCGCCATTTGCCAC 545
DB 433 TACAGCAGTGTCTCATCTCGGCTTCTATCAGTCTGAGCCGCTACCTGGCCATCGTCCAC 492
QY 546 GCACCAACAGTCAAGGCGCAAGAAAGTGTGGCTGAAAGGAGTCTATGTGGGCGTC 605
DB 493 GCCACCAACAGTCAAGGCGCAAGAAAGTGTGGCTGAAAGGAGTCTATGTGGGCGTC 552
QY 606 TGGATCCAGCCCTCTCTCTGACTATACCTGACTATATCTTTCGCGAGCTGAGCCAGGG 665
DB 553 TGGATCCCTGCTCTCTGACTATTCGCGACTTCTCTTGGCAACGCTCAGTGAGGCA 612
QY 666 GACATCAG 673
DB 613 GATGACAG 620

Search completed: May 17, 2004, 22:14:06
Job time : 337.298 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:50:31 ; Search time 1977.17 seconds
(without alignments)
10345.919 Million cell updates/sec

Title: US-09-367-052-5
Perfect score: 685
Sequence: 1 ccatactaatacagactcact.....gacatcagtcaggggatga 685

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_hic:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: em_gss_hum:**
- 18: em_gss_inv:**
- 19: em_gss_pln:**
- 20: em_gss_vrt:**
- 21: em_gss_fun:**
- 22: em_gss_mam:**
- 23: em_gss_mus:**
- 24: em_gss_pro:**
- 25: em_gss_rod:**
- 26: em_gss_pbg:**
- 27: em_gss_vrl:**
- 28: gb_gss1:**
- 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	608.8	88.9	906	13	BU523025
2	606.4	88.5	958	12	BG174412
3	590.0	86.1	753	12	BG315636
4	576.2	84.1	628	10	AW227957

5	544	79.4	955	12	BG173867
6	536.6	78.3	581	14	CF540844
7	511.4	74.7	585	10	BE627479
8	503	73.4	504	13	BX513842
9	487.8	71.2	886	10	BF101953
10	472.4	69.0	514	12	BG093377
11	470.8	68.7	765	12	BG387369
12	449.8	65.7	1201	13	BX418530
13	449.2	65.6	703	12	BI765768
14	446.8	65.2	802	14	CB988971
15	446.6	65.2	928	14	CD389217
16	446.6	65.2	1201	13	BX358346
17	445.8	65.1	730	13	BX421959
18	445.6	65.1	1034	12	BM920800
19	445.6	65.1	630	14	CB215245
20	445.6	65.1	702	12	BI761118
21	445.6	65.1	910	12	BI762229
22	445.6	65.1	913	12	BM051973
23	445.6	65.1	950	13	BQ718617
24	445.2	65.0	658	12	BI835125
25	445.2	65.0	742	12	BI597875
26	444.8	64.9	728	9	AU117058
27	444.6	64.9	952	13	BX362477
28	444.4	64.9	1201	13	BX353253
29	444.2	64.8	945	13	BX443258
30	444.2	64.8	976	13	BUI46456
31	444	64.8	746	12	BI917014
32	443.4	64.7	1201	13	BX462182
33	443	64.7	876	14	CA454255
34	442.4	64.6	1201	13	BX421287
35	442.2	64.6	1201	13	BX462627
36	442	64.5	782	12	BG685901
37	441.6	64.5	837	14	CB987769
38	440.6	64.3	905	12	BI754094
39	440	64.2	1051	12	BM452529
40	439.4	64.1	1201	13	BX421511
41	438.6	64.0	911	13	BX443334
42	437.6	63.9	875	12	BI756157
43	437.2	63.8	849	13	BU945383
44	436.6	63.7	1201	13	BX396365
45	436.4	63.7	1091	13	BX397685

ALIGNMENTS

RESULT 1
BU523025
LOCUS BU523025 906 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT 10154335 NCI_CGAP_C024 Mus musculus cDNA clone
IMAGE:6525663 5', mRNA sequence.
ACCESSION BU523025
VERSION BU523025.1 GI:22833475
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 906)
NTH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM14130 row: c column: 07
High quality sequence stop: 668.

SUMMARIES

FEATURES	Location/Qualifiers	source
ORIGIN	<p>Query Match 88.9%; Score 608.8; DB 13; Length 906; Best Local Similarity 98.7%; Pred. No. 8.6e-137; Matches 626; Conservative 0; Mismatches 2; Indels 6; Gaps 1;</p>	
	<p>52 AGGTACAGTGCCTCTCAGGCGTTTGTTGCTCCGGTAACACACACGCGGTAGACGA 111 1 ATGTACAGTGCCTCTCAGGCGTTTGTTGCTCCGGTAACACACGCGGTAGACGA 60</p>	
	<p>112 GTGTTCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAAGAAG 171 61 GTGTTCCATGGAACCGATC-----AGTATATACACTTCTGATAACTACTCTGAAGAAG 114</p>	
	<p>172 TGGGGTCTGGAGACTATGACTCCACACAGGAACCTCTTCCGGATGAACAGTCCATT 231 115 TGGGGTCTGGAGACTATGACTCCACACAGGAACCTCTTCCGGATGAACAGTCCATT 174</p>	
	<p>232 TCAATAGGACTTCTCTGCCACCATCTACTTTCATCATCTTCTTGACTGGCATAGTCGGCA 291 175 TCAATAGGACTTCTCTGCCACCATCTACTTTCATCATCTTCTTGACTGGCATAGTCGGCA 234</p>	
	<p>292 ATGGATGTGTATCTCTGTGTTACAGAGAAAGCTAAGGACGATGACGACAAGT 351 235 ATGGATGTGTATCTCTGTGTTACAGAGAAAGCTAAGGACGATGACGACAAGT 294</p>	
	<p>352 ACCGGTGCACCTGTCAAGTGGTGCACCTCTCTTTGTATCATCACATCCCTCTTGGGCAG 411 295 ACCGGTGCACCTGTCAAGTGGTGCACCTCTCTTTGTATCATCACATCCCTCTTGGGCAG 354</p>	
	<p>412 TTGATGCCATGGCTGACTGGTACTTTTGGAAATTTTGTGAAGCTGTCCATATCATCT 471 355 TTGATGCCATGGCTGACTGGTACTTTTGGAAATTTTGTGAAGCTGTCCATATCATCT 414</p>	
	<p>472 ACACTGTCAACTCTACACGCGTTCTATCTCGCTTCTATCAGCTGACCGGTACC 531 415 ACACTGTCAACTCTACACGCGTTCTATCTCGCTTCTATCAGCTGACCGGTACC 474</p>	
	<p>532 TCGCCATTGTCCACGCCACCAACAGTCAAAAGGCCAAGGAACTGTGCTGAAAGGCAG 591 475 TCGCCATTGTCCACGCCACCAACAGTCAAAAGGCCAAGGAACTGTGCTGAAAGGCAG 534</p>	
	<p>592 TCTATGTGGCGTCTGGATCCAGCGCTCTCTGACTATACCTGACTTATCTTTGCCG 651 535 TCTATGTGGCGTCTGGATCCAGCGCTCTCTGACTATACCTGACTTATCTTTGCCG 594</p>	
	<p>652 ACCTCAGCCAGGGGACATCAGTCAGGGGGATGA 685 595 ACCTCAGCCAGGGGACATCAGTCAGGGGGATGA 628</p>	
RESULT 2		
LOCUS	<p>BG174412 958 bp mRNA linear EST 06-FEB-2001</p>	
DEFINITION	<p>602334232F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457694 5', mRNA sequence.</p>	
ACCESSION	<p>BG174412</p>	
VERSION	<p>BG174412.1 GI:12681115</p>	
KEYWORDS	<p>EST.</p>	
SOURCE	<p>Mus musculus (house mouse)</p>	

480	GCATTGTCACGCCACGACGATCAAGCCAGGAACTGCTGGCTGAAAGCGATC	539
Db		
594	TATGGGGGCTGTGATCCAGGCCCTCCTCCTGACTATACCTGACTTCATCTTTGCCGAC	653
Qy		
540	TATGGGGGCTGTGATCCAGGCCCTCCTCCTGA-TATACCTGACTTCATCTTTGCCGAC	598
Db		
654	GTACGCCAGGGGGACATCAGTCAGGGGGATGA	685
Qy		
599	GTACGGCAGGGGGACATCAGTCAGGGGGATGA	630
Db		

RESULT 3
BG915636
LOCUS
DEFINITION
753 bp mRNA linear
BG915636
602814106F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4936465 5',
mRNA sequence.

ACCESSION	BG915636
VERSION	BG915636.1
KEYWORDS	GI:14296112
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	1 (bases 1 to 753)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: WGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LRAM10868 row: 1 column: 02
High quality sequence stop: 683.

```

FEATURES
source
Location/Qualifiers
1. 753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4936465"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: NotI. Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furch,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
```

Query Match	86.1%;	Score 590;	DB 12;	Length 753;
Best Local Similarity	98.7%;	Pred. No. 2.9e-132;		
Matches 626;	Conservative 0;	Mismatches 5;	Indels 3;	Gaps 3;
QY	50	GCAGGTAGCAGTGCACCTCTCGAGCGCTTGGTGTCTCCGGTAACCAACCAACCGGCTGTAGAGC	109	
DDb	10	GGAGGTAGCAGTGCACCTCTCGAGCGCTTGGTGTCTCCGGTAACCAACCAACCGGCTGTAGAGC	69	
QY	110	GAGTGTGGCCATGGAAACCGCATCAGTGTGAGTATATACACTTCTGATAACTACTCTGGAAGA	169	
DDb	70	GAGTGTGGCCATGGAAACCGCATCAGTGTGAGTATATACACTTCTGATAACTACTCTGGAAGA	139	
QY	170	AGTGGGTCTTGGAGACTATGACTCCCAACGAAGAACCTTGCTTCGGGATGAAAAACGTCCA	229	
DDb	130	AGTGGGTCTTGGAGACTATGACTCCCAACGAAGAACCTTGCTTCGGGATGAAAAACGTCCA	189	

```

/tissue types="tumor, gross tissue"
/dev stage="5 months"
/lab host="DH10B"
/clone lib="NCI CGAP Mam4"
/notes="Organ: mammary; Vector: pCMV
       Site 2: NotI; Cloned unidirectional
       library constructed by life techno
       providing samples: Lothar Hennighaus
       NIH Reference for transgenic model
       Differentiation 7, 3-11 (1996)

```

ORIGIN

Query March	84.1%	Score 576.2	DB 10	Length 628
Best Local Similarity	96.7%	Mid. No. 5.8e-129		
Matches 609	Conservative 0	Mismatches 19	Indels 2	Gaps 2
53	GGTAGCAGTCACCCCTCTGAGGCGTTTGGTGCTCCGGTAAACACACACGCGCTGTAGAGCGAG	112		
1	GGTAGCAGTCACCCCTCTGAGGCGTTTGGTGCTCCGGTAAACACACACGCGCTGTAGAGCGAG	60		
113	TGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTCTGATACTCTGAAAGAGT	172		
61	TGTTGCCATGGAACCGATCAGTGTGAGTATATACACTTCTGATACTCTCTGAAAGAGT	120		
173	GGGCTCTGGAGACTATGACTCCAAACAGGAACCCCTCTCCGGGATGAAAAACGTCCTCATTT	233		
121	GGGCTCTGGAGACTATGACTCCAAACAGGAACCCCTCTCCGGGATGAAACGTCCTCATTT	180		
233	CAATAGGATCTTCCTGCCCAACCATCTACTTCATCTCTTCTGACTGGGATAGTCGGCAA	292		
181	CAATAGGATCTTCCTGCCCAACCATCTACTTCATCTCTTCTGACTGGGATAGTCGGCAA	240		
293	TGGATTGGTGATCTGTGCTATGGTTTACACAGAGAAGCTTAAGAGCATGACGGAACAAGTA	352		
241	TGGATTGGTGATCTGTGCTATGGTTTACACAGAGAAGCTTAAGAGCATGACGGAACAAGTA	300		
353	CCGGCTGCACTGTCACTGGCTGACTCTCTTTTGTTCATCAACATCCCTCTCTGGGCAGT	412		
301	CCGGCTGCACTGTCACTGGCTGACTCTCTTTTGTTCATCAACATCCCTCTCTGGGCAGT	360		
413	TGATGCCATGCTGACTGTGCTATTTTGGGAATTTTGTGTAAAGGCTGTCCATCATCATCTA	472		
361	TGATGCCATGCTGACTGTGCTATTTTGGGAATTTTGTGTAAAGGCTGTCCATCATCATCTA	420		
473	CACCTGTCAACCTCTACAGCAGCGTTTCTCATCTCTGGCCTTCATCAGCCTGGACCGGTACCT	532		
421	CACCTGTCAACCTCTACAGCAGCGTTTCTCATCTCTGGCCTTCATCAACCTGGACCGGTACCT	480		
533	CGCCATTGTCAACGCCACCAACAGTCAAAGGCCAAGGAACTGCTGGCTGAAAAGGCAAGT	592		
481	-GCCATTGTCCATGCGCACCAACAGTCAAATGCCAAAAGAACTGCTGGCTG-AAAAGGCAAGT	538		
593	CTATGTGGGCGTCTGGATCCAGCGCTCTCTGACTATACCTTGACTTCATCTTTTGGCGGA	652		
539	CTATGTGGGCGTCTGGATCCAGCGCTCTCTGACTTCATCTTTACCTGACTTCATCTTTGGCGGA	598		
653	CCTCAGCCAGGGGACATCATGCTCAGGGGGA	682		
599	CCTCAGCCAGGGGACATCATTTTCAGGGGA	628		

RESULT 5	BG173867	955 bp	mRNA	linear	EST 06-FEB-2001			
LOCUS	60233928F1	NCI_CGAP_Mam1	Mus musculus	cDNA clone	IMAGE:4457119 5',			
DEFINITION	mRNA sequence.							
ACCESSION	BG173867							
VERSION	BG173867.1	GI:12680570						
KEYWORDS	EST.							
SOURCE	Mus musculus (house mouse)							
ORGANISM	Mus musculus							
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.							

```

REFERENCE 1 (bases 1 to 955)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10253 row: g column: 08
High quality sequence stop: 652.
Location/Qualifiers
1..955
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4457119"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
FEATURES
source

```

ORIGIN

Query Match	79.4%;	Score 544;	DB 12;	Length 955;
Best Local Similarity	95.6%;	Mismat. No. 4.6e-121;		
Matches 593;	Conservative	0;	Mismatches 20;	Indels 7; Gaps 3;
QY	73	GCCTTTGGTCTCCGCTAAACCAACCAACCGGCTGTAGAGCGAGTGTGGCCATGGAACCGATCA	132	
DB	1	GCCTTTGGTCTCCGCTAAACCAACCGGCTGTAGAGCGAGTGTGGCCATGGAACCGATCA	60	
QY	133	GTGTGAGTATATACACTTCTGTATACTCTCTGAAGAAGTGGGGTCTGGAGACTATGACT	192	
DB	61	GTGTGAGTATATACACTTCTGTATACTCTCTGAAGAAGTGGGGTCTGGAGACTATGACT	120	
QY	193	CCACCAAGGAACCGCTCTCCGGATGAAAACGCTCCATTCCAATAGGATCTTCTCTGCCCA	252	
DB	121	CCACCAAGGAACCGCTCTCCGGATGAAAACGCTCCATTCCAATAGGATCTTCTCTGCCCA	180	
QY	253	CCATCTACTTTCATCATCTCTTGTACTGGCATAGTCGGCAATGGAATGTGTGATCTCTGGTCA	312	
DB	181	CCATCTACTTTCATCATCTCTTGTACTGGCATAGTCGGCAATGGAATGTGTGATCTCTGGTCA	240	
QY	313	TGGGTTACCAAGAAGAAGCTTAAGAGCATGACCGACAGTACCGGCTGCACCTGTCACTGG	372	
DB	241	TGGGTTACCAAGAAGAAGCTTAAGAGCATGACCGACAGTACCGGCTGCACCTGTCACTGG	300	
QY	373	CTGACCTCTCTTGTTCATCACACTCCCCTTCTGGCGAGTTGATGCCATGGCTGACTGGT	432	
DB	301	CTGACCTCTCTTGTTCATCACACTCCCCTTCTGGCGAGTTGATGCCATGGCTGACTGGT	360	
QY	433	-ACTTTGGGAAATTTTGTGTAAAGGCTGCCATATCATCTACAC-----TGTCAACCTCT	486	
DB	361	AACTTTGGGAAATTTTGTGGTTAGGCTGGTCACTTTTATATACATGTGTAAACTCTCAT	420	
QY	487	ACAGAGCGTTCATCTCGGCTTCATCAGCTCTGACCGGTACCTCCGCCATGTGCCACG	546	
DB	421	AGAGAAGTGTTCATCTCTGGCTTCATCAGGCTCGAACCGGTACCTTCGCCATTTGTCCACG	480	
QY	547	CCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGACAGTCTATGTGGCGCTCT	606	
DB	481	CCACCAACAGTCAAAGGCCAAGGAAACTGCTGGCTGAAAAGGACAGTCTATGTGGCGCTCT	540	
QY	607	GGATCCCAAGCCCTCTCTCTGACTATACCTGACTCTCATCTTTTCCGA-CGTAGCCAGGGG	665	

Db 541 GGATCCAGCCCTCTCTGACTATACCTGACTTCATCTTGGCAGCCGTCAGCCAGGGG 600

Qy 666 GACATAGTCAGGGGGATGA 685

Db 601 GACATCAGTCAGGGGGATGA 620

RESULT 6

CF540844

LOCUS UI-M-GWO-cio-1-17-0-UI.r1 NIH_BMAP_GWO Mus musculus cDNA clone

DEFINITION IMAGE:30542392 5', mRNA sequence.

ACCESSION CF540844

VERSION CF540844.1 GI:34593367

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 581)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1..581
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30542392"
/tissue_type="whole eye"
/dev_stage="embryo 15.5, 16.5, 17.5, 18.5 dpc"
/lab_host="DH10B (11 phage resistant)"
/clone_lib="NIH_BMAP_GWO"
/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CTGGTCTCTC. This library was created for the University of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 78.3%; Score 536.6; DB 14; Length 581;

Best Local Similarity 98.6%; Pred No. 2.3e-119;

Matches 550; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 128 GATCAGTGTGAGTATACACTCTGATAACTCTCTGAAGAAAGTGGGGTCTGGAGACTA 187

Db 1 GATCAGTGTGAGTATACACTCTGATAACTCTCTGAAGAAAGTGGGGTCTGGAGACTA 60

Qy 188 TGATCCACAGAACCCCTCTCCGGGATGAAGCTCCATTCATAGGATCTTCCT 247

Db 61 TGATCCACAGGAACCCCTCTCCGGGATGAAGCTCCATTCATAGGATCTTCCT 120

Qy 248 GCCCACCATCTACTTTCATCATCTCTTCTTGAATGGCATAGTCGGCAATGGATTGGTGATCCT 307

Db 121 GCCCACCATCTACTTTCATCATCTCTTCTTGAATGGCATAGTCGGCAATGGATTGGTGATCCT 180

Qy 308 GGTATGCGTTTACCAGAAAGAGCTAAGGAGCATGACGGAACAAGTACCAGGCTGCACCTGTC 367

Db 181 GGTATGCGTTTACCAGAAAGAGCTAAGGAGCATGACGGAACAAGTACCAGGCTGCACCTGTC 240

Qy 368 AGTGGCTGACCTCTCTTCTTTCATCATCACTCCCTCTTGGGAGTGTGATGCCATGGGTGA 427

Db 241 AGTGGCTGACCTCTCTTCTTTCATCATCACTCCCTCTTGGGAGTGTGATGCCATGGGTGA 300

Qy 428 CTGGTACTCTTTGGGAAATTTTGTGTAAAGGCTGTCCATATCATCTACACCTGTCACACCTCTA 487

Db 301 CTGGTACTCTTTGGGAAATTTTGTGTAAAGGCTGTCCATATCATCTACACCTGTCACACCTCTA 360

Qy 488 CAGACGCTTCTCATCTCTTCTTTCATCATCACTCCCTCTTGGGAGTGTGATGCCATGGGTGA 547

Db 361 CAGACGCTTCTCATCTCTTCTTTCATCATCACTCCCTCTTGGGAGTGTGATGCCATGGGTGA 420

Qy 548 CACCAACAGTCAAAAGGCAAGGAAAGTGTCTGGCTGAAAGGCAAGTGTATGTGGGGCTGTG 607

Db 421 CACCAACAGTCAAGAGCCAGGAAAGTGTCTGGCTGAAAGGCAAGTGTATGTGGGGCTGTG 480

Qy 608 GATCCAGCCCTCTCTCTGACTATACCTGACTTCACTTTTCCGACGTCAGCCAGGGGGA 667

Db 481 GAT-CCAGCCCTCTCTCTGACTATACCTGACTTCACTTTTNGCGACGTCAGCCAGGGGGA 539

Qy 668 CATCAGTCAGGGGGATGA 685

Db 540 CATCAGTCAGGGGGATGA 557

RESULT 7

BE627479

LOCUS BE627479

DEFINITION BE627479.1 585 bp mRNA linear EST 24-AUG-2000

un52b07.y1 Soares thymus_2NbMT Mus musculus cDNA clone

IMAGE:3375541 5', similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION BE627479

VERSION BE627479.1 GI:9907901

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 585)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index.

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1085145
Seq primer: -40RP from Gibco
High quality sequence stop: 463.
Location/Qualifiers
1..585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3375541"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/clone_lib="Soares thymus_2NbMT"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified

RESULT 9

BF101953 886 bp mRNA linear EST 19-OCT-2000
LOCUS 601752919F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3980525 5',
mRNA sequence.
BF101953
ACCESSION BF101953.1 GI:10884479
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9176 row: e column: 06
High quality sequence stop: 665.
Location/Qualifiers
1..886
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3980525"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
FEATURES
source
Query Match 71.2%; Score 487.8; DB 10; Length 886;
Best Local Similarity 98.7%; Pred. No. 1.9e-107;
Matches 534; Conservative 0; Mismatches 2; Indels 5; Gaps 4;
QY 77 TTGGTGTCCGGTAACCAACACCGCTGTAGAGCGAGTGTGCCATGGAACCGATCAGTGT 136
DB 7 TTGGTGTCCGGTAACCAACACCGCTGTAGAGCGAGTGTGCCATGGAACCGATCAGTGT 66
QY 137 GAGTATATACACTTCTGATAACTCTGAAAGAGTGGGCTCGGAGACTAGTACTCAA 196
DB 67 GAGTATATACACTTCTGATAACTCTGAAAGAGTGGGCTCGGAGACTAGTACTCAA 126
QY 197 CAAGGAACCTGCTCCGGGATGAAGAGTCCATTTCATAGGATCTTCCTGCCACCAT 256
DB 127 CAAGGAACCTGCTCCGGGATGAAGAGTCCCA-TTCATAGGATCTTCCTGCCACCAT 185
QY 257 CTACTTTCATCATCTTCTTGACTGGCATAGTGGCAATGGATTGGTATCTCTGGTCATGGG 316
DB 186 CTACTTTCATCATCTTCTTGACTGGCATAGTGGCAATGGATTGGTATCTCTGGTCATGGG 245
QY 317 TTACAGAAGAGCTAAGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGCTGA 376
DB 246 TTACAGAAGAGCTAAGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGCTGA 305
QY 377 CCTCTCTTTGTCATCACATCTCCCTTCTGGGCGAGTTGATGCCATGGCTGATGGTACTT 436
DB 306 CCTCTCTC-TTGTTCATCACATCTCCCTTCTGGGCGAGTTGATGCCATGGCTGATGGTACTT 364
QY 437 TGGGAAATTTTGTGTAGGCTGTCCATATATCTACACTGTCAACCTCTCACAGCGGT 495

365 --GGGAAATTTTGTGTAGGCTGTCCATATCATCTACATGTCAACCTCTACAGCAGGT 422
QY 497 TCTCATCTCGCTTTCATCAGCTGAGCGGTACCTCGCCATGTCTCCACGCCACCAACAG 556
DB 423 TCTCATCTCGCTTTCATCAGCTGAGCGGTACCTCGCCATGTCTCCACGCCACCAACAG 482
QY 557 TCAAAGGCCAAGAACTGTGCTGAAAGAGCGAGTCTATGTGGGCGCTCTGGATCCAGC 616
DB 483 TCAAAGGCCAAGAACTGTGCTGAAAGAGCGAGTCTATGTGGGCGCTCTGGATCCAGC 541
QY 617 C 617
DB 542 C 542
RESULT 10
BG093377 514 bp mRNA linear EST 26-JAN-2001
LOCUS uc70904.y1 Soares mouse NMGB.bcell Mus musculus cDNA clone
DEFINITION IMAGE:333318 5' similar to SW:CCR4_MOUSE P70658 C-X-C CHEMOKINE
RECEPTOR TYPE 4 ; mRNA sequence.
ACCESSION BG093377
VERSION BG093377.1 GI:12575940
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 514)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1077482
Trace considered overall poor quality
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..514
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:333318"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares mouse NMGB.bcell"
/note="Organ: germinal B-cell; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTCACATCTGAAAGTGGAGCGCGCTGTTTCTTTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized; constructed by Bento Soares and
M. Fatima Bonaldo."

ORIGIN
Query Match 69.0%; Score 472.4; DB 12; Length 514;
Best Local Similarity 94.9%; Pred. No. 8e-104;
Matches 488; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 122 GGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGAGAGTGGGCTCTGG 181
DB 1 GGAACCGATCAGTGTGAGTATATACACTTCTGATAACTACTCTGAGAGAGTGGGCTCTGG 60
QY 182 AGACTATGATCTCAACAGAGAACCTTGTCTCGGGATGAAACGTCCTCATATAGGAT 241
DB 61 AGACTATGATCTCAACAGAGAACCTTGTCTCGGGATGAAACGTCCTCATATAGGAT 120

/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-CNI"
/polylinker="pTT73D-Pac (Pharmacia) with a modified
Not I; Site 2: Eco RI; The UI-R-CNI
library is a subtracted library derived from the following
pool of seven normalized rat libraries: normalized rat
seminal vesicles, normalized rat penis, normalized rat
bladder, normalized rat cervix, normalized rat brown
adipose, normalized rat fundus, and normalized rat
salivary gland. It was constructed according to the
procedure described by Bonaldo, Lennon & Soares (Genome
Research, Volume 6: 791-806, 1996). For construction of
the CNI library, plasmid DNA from the pool of normalized
libraries was electroporated into competent bacteria for
the production of single-stranded circular DNA. This was
then used as a tracer in a subtractive hybridization with
a driver (PCR amplified inserts from a plasmid DNA template
preparation) comprising: a) a pool of about 34,000 clones
from the Rat Unigene Set corresponding to plates R-5-AA-NN
excluding plates R-5-MM and MN. This pool represented 40%
of the final driver population. b) a pool of about 29,000
clones from subtracted libraries CA0 and CA1 corresponding
to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AXZ through
R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS,
R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through
R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLV through R-CA0-BMA,
R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through
R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA0-BOB through
R-CA1-BDA, R-CA1-BHZ through R-CA1-BJF, R-CA1-BJR,
R-CA1-BJT through R-CA1-BKS, R-CA1-BKD, R-CA1-BKF,
R-CA1-BKI, R-CA1-BKT, R-CA1-BJF, R-CA1-BLH through
R-CA1-BLN, R-CA1-BLN, R-CA1-BLU-V, R-CA1-BNR, and
R-CA1-BLE. The resulting pool represented 20% of the
final driver population. c) a pool of about 15,000 clones
from non-normalized libraries CS08, CT08, CU08, CW08, CX08
and normalized libraries CS0, CT0, CU0, CW0, and CX0
corresponding to plates R-CS08-CBD through R-CS08-CBO,
R-CT08-CAM through R-CT08-CAX, R-CU08-CBP through
R-CU08-CCA, R-CW08-CCB through R-CW08-CCM, R-CX08-CCN
through R-CX08-CCX, R-CS0-BSD, R-CS0-BTD through
R-CS0-BTV, R-CS0-BVM, R-CT08-BTW through R-CT0-EUP,
R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through
R-CW0-BWP, R-CX0-BXN through R-CX0-BXO, R-CX0-BWQ through
R-CX0-BXM. The resulting pool represented 5% of the final
driver population. d) a pool of about 5,000 clones
from non-normalized eye library CV0 and 4,000 clones
normalized eye library CV1) corresponding to plates
R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC,
R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through
R-CV1-BVU. This pool represented about 5% of the final
driver population. e) A pool of about 10,000 clones from
subtracted library BS2, BV0 and BV0P (7-9.5 kb cDNA
library fraction from rat whole embryo) and BX0 (0.5-7kb
cDNA library fraction from rat whole embryo) corresponding
to plates R-BS2-BPB through R-BS2-BFP, R-BV0-BNK through
R-BV0-ANR, R-BV0-AOI through R-BV0-AOK, and R-BX0-AQY
through R-BX0-ASH. The resulting pool represented 5% of
the final driver population. f) a pool of about 7,000
clones from the seven non-normalized libraries that make
up the tracer including CY0, CZ0, DA0, DB0, DC0, DD0, and
DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ,
R-CZ0-BYA through R-CZ0-BYI, R-DC0-BZB-C, R-DA0-BYU
through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ
through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY
through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA,
R-DD0-CBB-C, and R-DB0-CAB through R-DE0-CAL. The
resulting pool represented about 10% of the final driver
population. g) a pool of about 2,000 clones from the pool
of normalized libraries, CNO, that makes up the tracer.
The corresponding plates are R-CNO-BKW through R-CNO-BLD,
R-CNO-BLE, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT,
R-CNO-BLW-X, R-CNO-BME, and R-CNO-BMF through R-CNO-BML.

This pool represented 5% of the final driver population.
h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bkx-a-05-0-UI, bkx-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, bla-h-12-0-UI, bla-f-02-0-UI, bla-a-11-0-UI, bla-c-12-0-UI, bla-d-1-08-0-UI, bla-d-02-0-UI, bla-q-h-04-0-UI, bla-r-a-05-0-UI, bla-r-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG TISSUE=cervix
TAG LIB=UI-R-CNI
TAG_SEQ=GRACCA"

ORIGIN	
Query Match	68.7%; Score 470.8; DB 12; Length 765;
Best Local Similarity	92.2%; Pred. No. 2.4e-103;
Matches	496; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY	136 TGAGTATATACACTCTGTAAGTACTCTGAGAGTGGGGCTCGAGACTATGCTCCA 195
Db	105 TGGAAATATACACTTCGGATAATCTCCGAAAGTAGGGTCTGAGACTATGCTCCA 164
QY	196 ACAAGGAACCTCTCTCCGGATGAAAGCTCCATTTCAATAGGATCTTCCTGCCACCA 255
Db	165 ACAAGGAACCTCTCTCCGGATGAAAGCTTCACAGGATCTTCCTGCCACCA 224
QY	256 TCTACTTCATCATCTCTTGACTGGCATAGTCGGCAATGGATGGTGGTCTGTCATGG 315
Db	225 TCTATTTTATATCTCTTGACTGGCATAGTCGGCAATGGTGGTGGTGGTGGTGGT 284
QY	316 GTTACAGAGAGAGCTTAAGGAGCATCAGGACCAAGTACCGGCTGCACCTGTGAGTGGCTG 375
Db	285 GTTACAGAGAGAGCTGAGGAGCATCAGGACCAAGTACCGGCTGCACCTGTGAGTGGCTG 344
QY	376 ACTCTCTTTGTCATCACACTCCCTTTCTGGGCACTGATGCGCATGGTGTGACTGCTACT 435
Db	345 ACTCTCTTTGTCATCACACTCCCTTTCTGGGCACTGATGCGCATGGTGTGACTGCTACT 404
QY	436 TTGGGAAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCAACCTTACAGCAGG 495
Db	405 TTGGGAAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCAACCTTACAGCAGG 464
QY	496 TTCTCATCTCTGGCTTCATCAGCCTGAGACCGGTACCTCGCCATTTGCCACGCCACCA 555
Db	465 TTCTCATCTCTGGCTTCATCAGCCTGAGACCGGTACCTCGCCATTTGCCACGCCACCA 524
QY	556 GTCAAGGCCCAAGAACTGCTGGCTGAAAGGAGTCTATGTGGCGTCTGATCCCG 615
Db	525 GCCAGAGGCCGAGGAGTCTGCTGGCTGAAAGGAGTCTATGTGGGTGTCTGGATCCCCG 584
QY	616 CCTCTCTCTGACTATACCTGACTTCTATCTTTCGCGAGCTCAGCAGGGGGGACATCAG 673
Db	585 CCTCTCTCTGACTATACCTGACTTCTATCTTTCGCGAGCTCAGCAGGGGGGAGCGAG 642

RESULT 12
BX418530
LOCUS
DEFINITION BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF006VJ16 5-PRIME, mRNA sequence.
ACCESSION BX418530
VERSION BX418530.1
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber.C., Jessee,J. and Polayes.D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF006DE08QPI&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF006DE08QPI.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS0DF006VJ16"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

FEATURES
source

ORIGIN

Query Match 65.7%; Score 449.8; DB 13; Length 1201;
Best Local Similarity 82.8%; Pred. No. 3.6e-98;
Matches 526; Conservative 1; Mismatches 103; Indels 5; Gaps 1;
QY 39 CCCCAGGAGGTGACAGTACCTCTCTGAGGCGCTTGGTCTCCGGTAAACCCAC 98
Db 61 CCCCAGGAGGTGACAGTACCTCTCTGAGGCGCTTGGTCTCCGGTAAACCCAC 120
QY 99 GCGTGTGAGGAGTGTGGCATGGAACCGATCAGTGTGAGTATATACACTTCTGATAAC 158
Db 121 ATCTGGAGAACACCGGTTACCATGGAGGGA-----TCAGTATATACACTTCAGATAAC 175
QY 159 TACTCTGAAGAGTGGGCTCTGGAGACTATGACTCCCAAGAACCTCTCTCCGGGAT 218
Db 176 TACACCGAGGAATGGGCTCAGGGGACTATGATTCATGAGGAGAACCTCTGTTCCGTGAA 235
QY 219 GAAACAGTCCATTTCAATAGGATCTTCTGCCACCACTCTCTCATCTCTCTTCTGACT 278
Db 236 GAAATGCTAATTTCAATATAAATCTTCTGCCACCACTCTCTCATCTCTCTTAACT 295
QY 279 GGCATAGTCGGCATGATGGTGATCTCTGGTATGGGTTACCGAGAGAGCTAAGAGC 338
Db 296 GGCATTTGGGCAATGATGGTGTCTCTGAGTGGGCGACCTCTCTCTTGTGATCAGCTT 355
QY 339 ATGACGCAAGTACCGGCTGCACCTCTGAGTGGGCTGACCTCTCTTGTGATCAGACTC 398
Db 356 ATGACGCAAGTACAGGCTGCACCTCTGAGTGGGCGACCTCTCTTGTGATCAGCTT 415
QY 399 CCTTCTGGGCAATGATGGCATGGCTGATCTTTGGGAAATTTTGTGTAAGGCT 458
Db 416 CCTTCTGGGCAATGATGGCGGAAACTGGTACTTTGGGAACTTCTCTATGCAAGGCA 475
QY 459 GTCCATATCATCTACACTGTCAACCTCTACAGGAGCGTCTCTCATCTCGGCTTCTCAGC 518
Db 476 GTCCATGTATCTACAGCACTCACTCTACAGGAGTGTCTCTCATCTCGGCTTCTCAGT 535
QY 519 CTGACCGGTAACCTCGGCAATGTCACGCCACCAAGTCAAGGCCAAGAACTGCTG 578

Db	536	CTGACCGCTACCTGGCCATCGTCCACGCCACACAGTCAGAGGCCAAGGAGCTGTG	595
Qy	579	GCTGAAGAAGCAGTCTATGTGGGCGTGTGATCCAGCCCTCTCTGACTATACCTGAC	638
Db	596	GCTGAAGAAGGTGCTATGTGTGGCGTGTGATCCCTGCTGCTACTATCCCGAC	655
Qy	639	TTCATCTTTTCCGACGTCAGCAGGGGGACATCAG	673
Db	656	TTCATCTTTGGCAACGTGACGTCAGTACGACAGTACAG	690

RESULT 13	BI765768	LOCUS	BI765768	703 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603046522F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186837 5', mRNA sequence.						

ACCESSION	E178576B
VERSION	B178576B.1
KEYWORDS	GI:15757346
EST.	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 703)
TITLE	NIH-MGC http://mgc.ncl.nih.gov/
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LLNL at:

```

http://image.llnl.gov
Plate: LLAM11466 row: 1 column: 06
High quality sequence start: 24
High quality sequence stop: 703.
Location/Qualifiers
FEATURES

```

```

1. .703
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5186837"
/lab_host="DH10B"
/clone_lib="NIH MGC 116"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is a
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."

```

ORIGIN

Query Match	65.6%;	Score 449.2;	DB 12;	Length 703;
Best Local Similarity	81.8%;	Pred. No. 3.9e-98;		
Matches 532;	Conservative	0;	Mismatches 113;	Indels 5;
				Gaps 1;

24	GGGCTCAGCGGCGCCCGGGCAGGTGCAGGTAGCAGTCAACCTCTGAGCGCTTTGGTGC	83
	DB	
9	GGTCCCGAGTCGTTGATCTCGGCACGAGTAGCAAGTGACCGCAGGGCTGAGTGC	68
84	TCGGTAAACCAACACCGGCTGTAGAGCGAGTGTTCATGGAAACCGATCAGTGTGAGTATA	143
	QY	
69	TCAGTAGCACCGCATCTGGAGAACACGCGGTACCATGGAGGGA-----TCAGTATA	123
	DB	
144	TACACTCTTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAGGAA	203
	QY	

Db	124	TACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGACTATGACTCCATGAGGAA	183
QY	204	CCCTGCTTCGGGATGAAACGTCATTCAATAGGATCTTCTCTGCCACCATCTACTTC	263
Db	184	CCCTGTTTCGGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCACCATCTACTCC	243
QY	264	ATCATCTTCTGACTGGCATAGTCGGCAATGGATTTGGTGATCTCTGCTCATGGGTACCAG	323
Db	244	ATCATCTTCTTAAGTGGCATTTGGGCAATGGATTGGTTCATCTCTGCTCATGGGTACCAG	303
QY	324	AGAAGCTAAGGAGCATGACGACAAGTACCGCTGCACCTGTCACTGGTGAACCTCCTC	383
Db	304	AGAAACTGAGAAGCATGACGACAAGTACAGGCTGCACCTGTCACTGGCCGACCTCCTC	363
QY	384	TTTGTTCATCACACTCCCTTCTGGGCAGTTGATGCCATGGCTGACTGCTACTTTGGGAAA	443
Db	364	TTTGTTCATCAGTTCCCTTCTGGGCGATTTGATGCCGTGGCAACTGGTACTTTGGGAAC	423
QY	444	TTTTTGTGTAAGCTGTGCATATCATCTACACTGTCAACCTCTACAGCAGCTTCTCATC	503
Db	424	TTTCCTATGCAAGCGAGTCCATGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTCATC	483
QY	504	CTGGCCTTCATCAGCCTGACCGGTACCTCGGCATGTTCACGCCACCAACAGTCAAAGG	563
Db	484	CTGGCCTTCATCAGTCTGGACCGCTACCTGGCCATCGTCCACGCCACCAACAGTCAGAGG	543
QY	564	CCAAGGAAACTGCTGCTGCAAAAGGCAGTCTATGTGGGGCTCTGGATCCACAGCCCTCCTC	623
Db	544	CCAGGAAGCTGTGGCTGAAAGGTGGTCTATGTGGCGCTCTGGATCCCTGCCCTCTCG	603
QY	624	CTGACTATACCTGACTTCATCTTTGCCAGCTCAGCCAGGGGAGACATCAG	673
Db	604	CTGACTATTTCCGACTTCATCTTTGCCAAGTCACTGGAGGCAGATGACAG	653
RESULT 14			
CB988971			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

RESULT 14					
CB988971					
LOCUS	CB988971	802 bp	mRNA	linear	EST 01-MAY-2003
DEFINITION	AGENCOURT_13833109 NIH_MGC_147 Homo sapiens CDNA clone				
	IMAGE:30346124 5', mRNA sequence.				

ACCESSION	CB988971
VERSION	CB988971.1
KEYWORDS	GI:30283491
SOURCE	EST.
ORGANISM	Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 802)
NIH-MGC <http://mgc.ncbi.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov
Tissue procurement: Dr. Stefan Hansson
cDNA Library preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: NDAM385 row: b column: 21
High quality sequence stop: 663.

```

FEATURES
source
Location/Qualifiers
1. 802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30346134"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"

```

/notes="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3', size-selected for average insert size 2.3 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: This is a NIH_MGC library."

ORIGIN

Query Match 65.2%; Score 446.8; DB 14; Length 802;
Best Local Similarity 83.0%; Pred. No. 1.6e-97;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGGTGCAGGTAGCAGTACCTCTGAGCGCTTTCGTCCTCGGTAAACACACGCGGTG 103
DB 40 GCGGCAGCAGGTAGCAAGTACGCGCGGCTGAGTCTCCAGTAGCCACCGCATCTG 99
QY 104 TAGAGGAGTGTGCCATGGAACGATCAGTGTGATATATACATCTCTGATTAATCTTC 163
DB 100 GAGAACACGCGGTACCATGGA-----GGGGATCAGTATATACATCTTCAGATAACTACAC 154
QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCACAGGAACCTGCTTCCGGGATGAAAA 223
DB 155 CGAGGAATGGGCTCAGGGACTATGCTCCATGAAGGAACCTGTTCCGTGAAGAAA 214
QY 224 CPTCCATTTCAATAGGATCTTCTGCCACCATCTACTTTCATCATCTTCTTCACTGGCAT 283
DB 215 TGCTAATTTCAATAAATCTTCTGCCACCATCTACTCCATCATCTTCTTAACCTGGCAT 274
QY 284 AGTCGGCAATGGATTGGTGTCTCTGTCATGGTTACCAAGAGAGCTAAGGAGCATGAC 343
DB 275 TGTGGCAATGGATTGGTGTCTCTGTCATGGTTACCAAGAGAGCTAAGGAGCATGAC 334
QY 344 GGAACAAGTACCGCTGCACCTGTGTCAGTGGTGCACCTCTCTTGTTCATCACACTCCCTT 403
DB 335 GGAACAAGTACCGCTGCACCTGTGTCAGTGGTGCACCTCTCTTGTTCATCACACTCCCTT 394
QY 404 CTGGCAGTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
DB 395 CTGGCAGTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
QY 464 TATCATCTACATGTCACCTCTACAGCAGCTGCTCTCATCTGCTGCTGCTGCTGCTGCTGCT 523
DB 455 TGTCTATACACAGTCACTCTACAGCAGTGTCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCT 514
QY 524 CCGGTACCTGCGCATTTGTCAGCCACCAACAGTCAAGGCGCAAGAACTGCTGCTGCTGCTGCT 583
DB 515 CCGGTACCTGCGCATTTGTCAGCCACCAACAGTCAAGGCGCAAGAACTGCTGCTGCTGCTGCT 574
QY 584 AAGGAGTCTATGTGGGCTGCTGATCCAGCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
DB 575 AAGGAGTCTATGTGGGCTGCTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
QY 644 CTTTGGCAGCTCAGCGAGGGGACATCAG 673
DB 635 CTTTGGCAGCTCAGTGAGGAGATGACAG 664

RESULT 15

CD389217
LOCUS
DEFINITION AGENCOURT_14302495 NIH_MGC_173 Homo sapiens cdna 5', mRNA sequence.
ACCESSION CD389217
VERSION CD389217.1 GI:31227843
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 928)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE
JOURNAL

COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: csaps-re@mail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDKM74 row: 1 column: 08
High quality sequence stop: 797.
High quality sequence stop: 797.

FEATURES

source

1..928
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem
cells"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_173"
/note="vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR PRIMING - oligo dt; METHOD - full-length enriched;
LIBR PROVIDER - Bradfield"

ORIGIN

Query Match 65.2%; Score 446.8; DB 14; Length 928;
Best Local Similarity 83.0%; Pred. No. 1.7e-97;
Matches 523; Conservative 0; Mismatches 102; Indels 5; Gaps 1;
QY 44 GCAGGTGCAGTACGTCAGTACCTCTGAGCGCTTTCGTCCTCGGTAAACACACGCGGTG 103
DB 34 GCGGCAGCAGGTAGCAAGTACGCGCGGCTGAGTCTCCAGTAGCCACCGCATCTG 93
QY 104 TAGAGCAGTGTGCCATGGAACGATCAGTGTGATATATACATCTCTGATTAATCTTC 163
DB 94 GAGAACACGCGGTACCATGGA-----GGGGATCAGTATATACATCTTCAGATAACTACAC 148
QY 164 TGAAGAAGTGGGTCTGGAGACTATGACTCCACAGGAACCTGCTTCCGGGATGAAAA 223
DB 149 CGAGGAATGGGCTCAGGGAGCTATGATCTCATGAGGAACCTGTTTCGTGAAGAAA 208
QY 224 CGTCCATTTCAATAGGATCTTCTGCGCCACCATCTACTTCATCATCTCTTGTGACTGGCAT 283
DB 209 TGTCTAATTTCAATAAATCTTCTGCGCCACCATCTACTTCATCATCTCTTAACTGGCAT 268
QY 284 AGTCGCAATGGATTGGTGTCTCTGTCATGGTTACCAAGAGAGCTAAGGAGCATGAC 343
DB 269 TGTGGCAATGGATTGGTGTCTCTGTCATGGTTACCAAGAGAGCTAAGGAGCATGAC 328
QY 344 GGAACAAGTACCGCTGCACCTGTGTCAGTGGTGCACCTCTTGTTCATCATCTCCCTT 403
DB 329 GGAACAAGTACCGCTGCACCTGTGTCAGTGGTGCACCTCTTGTTCATCATCTCCCTT 388
QY 404 CTGGCAGTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 463
DB 389 CTGGCAGTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 448
QY 464 TATCATCTACACTGTCAACCTCTACAGCAGCTTCTCATCTCTGCGCTTCTATCAGCCTGGA 523
DB 449 TGTCTATCTACAGTCAACCTCTACAGCAGTGTCTCTCATCTCTGCGCTTCTATCAGTCTGA 508
QY 524 CCGGTACCTGCGCATTTGTCAGCCACCAACAGTCAAGGCGCAAGAACTGCTGCTGCTGCTGCT 583
DB 509 CCGGTACCTGCGCATTTGTCAGCCACCAACAGTCAAGGCGCAAGAACTGCTGCTGCTGCTGCT 568
QY 584 AAGGAGTCTATGTGGGCTGCTGATCCCGAGCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643

Db 569 AAAGTGGTCTATGTTGGCGTCTGGATCCCTGCCCCCTCTGCTGACTATTCCCGACTTCAT 628
QY 644 CTTTGGCGACGTCAGCCAGGGGGACATCAG 673
Db 629 CTTTGGCCAAACGTCAGTGAGGCAGATGACAG 658

Search completed: May 17, 2004, 16:09:07
Job time : 1983.17 secs

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 04:56:59 ; Search time 6852.38 Seconds
(without alignments)
10714.982 Million cell updates/sec

Title: US-09-367-052-7
Perfect score: 1694
Sequence: 1 atatacactctgataaata.....taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1694	100.0	1877	10	D87747	D87747 Mus musculus
2	1680.4	99.2	1817	10	BC031665	BC031665 Mus muscu
3	1599.8	94.4	175370	2	AC145861	AC145861 Pan ttrcgl
4	1599.8	94.4	185215	2	AC147556	AC147556 Mus muscu
5	1596.2	94.2	1809	10	AB000803	AB000803 Mouse mRN
6	1383.8	81.7	1575	10	MMLESTRPT	X99582 M.musculus
7	1383.8	81.7	3770	10	MMLESTRGN	X99581 M.musculus
8	1168.2	69.0	190149	2	AC122097	AC122097 Rattus no
9	1166.6	68.9	1223	10	MMU59760	U59760 Mus musculus
10	1151.8	68.0	1422	10	MMLCR13	280113 Mus musculus
11	1139.8	67.3	1180	10	MMLCR12	280112 Mus musculus
12	1114	65.8	240931	2	AC107097	AC107097 Rattus no
13	1054.2	62.2	3366	10	MMU65580	U65580 Mus musculus
14	946.4	55.9	1362	10	RNU54791	U54791 Rattus norv
15	933.4	55.1	1582	6	AX481810	AX481810 Sequence
16	933.4	55.1	1645	9	HSPNRYLA	X71635 H.sapiens m
17	933.4	55.1	1662	9	BC020968	BC020968 Homo sapi
18	933.4	55.1	1664	6	E37853	E37853 Anti-HIV ag
19	933.4	55.1	1664	6	AR222504	AR222504 Sequence
20	933.4	55.1	1664	6	AR380668	AR380668 Sequence
21	933.4	55.1	1664	9	HUMHM89	D10924 Homo sapien
22	933.4	55.1	1664	11	G28514	G28514 SHGC-31461
23	933.4	55.1	1667	9	AX129916	AX129916 Homo sapi
24	933.4	55.1	1679	6	AX287116	AX287116 Sequence
25	933.4	55.1	1679	6	AX548790	AX548790 Sequence
26	933.4	55.1	1679	6	AX705078	AX705078 Sequence
27	933.4	55.1	1737	6	AR015970	AR015970 Sequence
28	933.4	55.1	1737	6	AR060748	AR060748 Sequence
29	933.4	55.1	1737	6	AR070433	AR070433 Sequence
30	933.4	55.1	1737	6	AR103430	AR103430 Sequence
31	933.4	55.1	1737	6	II13753	II13753 Sequence 4
32	933.4	55.1	1737	6	I24455	I24455 Sequence 3
33	933.4	55.1	1944	6	E33806	E33806 Human splic
34	933.4	55.1	1950	9	AF147204	AF147204 Homo sapi
35	933.4	55.1	3733	9	HSCXCR4	Y14739 Homo sapien
36	933.4	55.1	5161	9	AF005058	AF005058 Homo sapi
37	933.4	55.1	172281	9	AC068492	AC068492 Homo sapi
38	931.8	55.0	8747	9	AF052572	AF052572 Homo sapi
39	931.8	55.0	8747	9	HS4224869	AJ224869 Homo sapi
40	931.4	55.0	1670	6	AX409496	AX409496 Sequence
41	931.4	55.0	1670	6	AX663641	AX663641 Sequence
42	931.4	55.0	1670	9	HUMGEPR	L06797 Human (Clon
43	908.6	53.6	1050	10	AF452185	AF452185 Rattus no
44	905.4	53.4	1050	10	RNT90610	U90610 Rattus norv
45	903.8	53.4	1637	9	HUMSTSR	M99293 Homo sapien

ALIGNMENTS

RESULT 1
D87747
LOCUS D87747 D87747 1877 bp mRNA linear ROD 07-FEB-1999
DEFINITION Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION D87747
VERSION D87747.1 GI:1772445
KEYWORDS murine CXCR-4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (sites)
AUTHORS Nagasawa,T., Nakajima,T., Tachibana,K., Iizasa,H., Bleul,C.C.,
Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
Kishimoto,T.

TITLE Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)

MEDLINE 9721456

REFERENCE 8962122

REFERENCE 2 (bases 1 to 1877)

AUTHORS Nagasawa,T.T.

TITLE Molecular cloning of murine PBSF/SDF-1 receptor

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 1877)

AUTHORS Nagasawa,T.

TITLE Direct Submission

JOURNAL Submitted (05-SEP-1996) Takashi Nagasawa, Research Institute, Osaka Medical Center, Department of Immunology, Murdoccho 840, Izumi, Osaka 590-02, Japan (E-mail:Immunoleosk.threeweibnet.or.jp, Tel:0725-56-1220, Fax:0725-57-3021)

FEATURES

Location/Qualifiers

1..1877

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/cell_line="DW34"

/cell_type="pre-B cell"

/tissue_type="bone marrow"

1..1877

/genes="murine CXCR-4"

120..1199

/gene="murine CXCR-4"

/function="PBSF/SDF-1 receptor"

/codon_start=1

/evidence="experimental"

/product="murine CXCR-4"

/protein_id="BAAL3451.1"

/db_xref="GI:1772446"

/translation="MEDISSIVTSNYSSEVSGDYDSNKEPCFRDENVHNRIFLP TIYPIELTGIVGNLIVLVYGVKLRSMTRVRLSVADLLFVITLFPWAVDAMA DWYKFLKRAVHLIYVSVLLAILSLRDLAIIVHATNSQPRKLLAEKAVYV GWIPLALLIPPEIPADVSQDIIISQDDRYICDRLYPSLWNVVQFQHMVGLLP GIVLSYCIILIIKLSHSGHQRKALKTIYVILIAFFACWLYFYVYISDFILGV IKQCDPESIVHWKISITAEALAFPHCLNPILYAFILGAPKPSAQHALNSMRGSLK ILSKRGEGGSHSVSTESSESPHSS"

source

gene

CDS

ORIGIN

Query Match 100.0%; Score 1694; DB 10; Length 1877;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAG 60

Db 141 ATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAG 200

Qy 61 GAACCTGCTCCGGGATGAAGCTCCATTCATAGGATCTCTCGCCCAACATCTAC 120

Db 201 GAACCTGCTCCGGGATGAAGCTCCATTCATAGGATCTCTCGCCCAACATCTAC 260

Qy 121 TTCAATCATCTTCTGATGCGCATAGTCGGAATGGATGGTGATCTCGGTGATGGGTAC 180

Db 261 TTCAATCATCTTCTGATGCGCATAGTCGGAATGGATGGTGATCTCGGTGATGGGTAC 320

Qy 181 CAGAGAAGCTTAGGACATGACGCAAGTACCGGCTGACCTGTCAGTGGCTGACCTC 240

Db 321 CAGAGAAGCTTAGGACATGACGCAAGTACCGGCTGACCTGTCAGTGGCTGACCTC 380

Qy 241 CTCCTTTGTCATCACATCCCTCTCTGGCGAGTGTGATGCCATGGCTGACTGGTACTTTGGG 300

Db 381 CTCCTTTGTCATCACATCCCTCTCTGGCGAGTGTGATGCCATGGCTGACTGGTACTTTGGG 440

Qy 301 AAATTTTGTAGGCTGTCCATATCATCTACATGTCAACTCTACAGAGCGTTCTC 360

Db 441 AAATTTTGTAGGCTGTCCATATCATCTACATGTCAACTCTACAGAGCGTTCTC 500

Qy 361 ATCTGGGCTTTCATCAGCTGAGCGGTACCTGGCATTGTCCAGCCACCAACAGTCAA 420

Db 501 ATCTGGGCTTTCATCAGCTGAGCGGTACCTGGCATTGTCCAGCCACCAACAGTCAA 560

Qy 421 AGGCCAAGGAAACTGCTGGCTGAAAGGAGTCTATGTGGGGCTGTGATCCAGCCCTC 480

Db 561 AGGCCAAGGAAACTGCTGGCTGAAAGGAGTCTATGTGGGGCTGTGATCCAGCCCTC 620

Qy 481 CTCCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCCAGGGGAGACATCAGTCAGGG 540

Db 621 CTCCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCCAGGGGAGACATCAGTCAGGG 680

Qy 541 GATGACGATACATCTGACCGGCTTTACCCCGATAGCCTGTGGATGGTGGTGTCAA 600

Db 681 GATGACGATACATCTGTCACCGGCTTTACCCCGATAGCCTGTGGATGGTGGTGTCAA 740

Qy 601 TTCCAGCAGATATAAGTGGTCTCATCTGCTGCCGAGTCGTCATCTCTCTCTCTCTCTCTCT 660

Db 741 TTCCAGCAGATATAAGTGGTCTCATCTGCTGCCGAGTCGTCATCTCTCTCTCTCTCTCTCT 800

Qy 661 ATCATCATCTCTAAGCTGTCTACCTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 720

Db 801 ATCATCATCTCTAAGCTGTCTACCTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 860

Qy 721 ACAGTCATCTCTAAGCTGTCTACCTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 780

Db 861 ACAGTCATCTCTAAGCTGTCTACCTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 920

Qy 781 ATCGACT 840

Db 921 ATCGACT 980

Qy 841 CACAAGTGGATCTCCATCACAAGGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

Db 981 CACAAGTGGATCTCCATCACAAGGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040

Qy 901 CTCTATGCT 960

Db 1041 CTCTATGCT 1100

Qy 961 AGCAGAGCTCTCAGCTCTCAGATCTCTTCCAAAGGAAAGCGGGGTGACACTCTCTCTCTCT 1020

Db 1101 AGCAGAGCTCTCAGCTCTCAGATCTCTTCCAAAGGAAAGCGGGGTGACACTCTCTCTCTCT 1160

Qy 1021 TCCAGGAGTCTCAGATCTCTCAGATCTTCACTCCAGCTAAACCTTATGCAAGACTTATAT 1080

Db 1161 TCCAGGAGTCTCAGATCTCTCAGATCTTCACTCCAGCTAAACCTTATGCAAGACTTATAT 1220

Qy 1081 AAT 1140

Db 1221 AAT 1280

Qy 1141 ACTGACCACTCTCTCAGATCT 1200

Db 1281 ACTGACCACTCTCTCAGATCT 1340

Qy 1201 TCTAGTCT 1260

Db 1341 TCTAGTCT 1400

Qy 1261 GAATGAGGCTCTAGCAGGACCTGTGGCAGGTTCTTAGTAGCTGTCTATCTGTGTGTAG 1320

Db 1401 GAATGAGGCTCTAGCAGGACCTGTGGCAGGTTCTTAGTAGCTGTCTATCTGTGTGTAG 1460

Qy 1321 GACTGTAGAACTGTAGAGGAAAGAACTGAACATTCACAGAAATGTGTGTAATTCGAATAAA 1380

Db 1461 GACTGTAGAACTGTAGAGGAAAGAACTGAACATTCACAGAAATGTGTGTAATTCGAATAAA 1520

Qy 1381 GCTAGCCGCTGATCTCAGCTGTCTGCTGATAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440

Db 1521 GCTAGCCGCTGATCTCAGCTGTCTGCTGATAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1580

Qy 1441 CCACCCCAACCCCAACCCCAATCT 1500


```
* * * 64667 64766: gap of unknown length
* * * 64767 65994: contig of 1228 bp in length
* * * 65995 66094: gap of unknown length
* * * 66095 67430: contig of 1336 bp in length
* * * 67431 67530: gap of unknown length
* * * 67531 68382: contig of 1452 bp in length
* * * 68383 68082: gap of unknown length
* * * 68083 70891: contig of 1809 bp in length
* * * 70892 70991: gap of unknown length
* * * 70992 72172: contig of 1181 bp in length
* * * 72173 72272: gap of unknown length
* * * 72273 73614: contig of 1342 bp in length
* * * 73615 73714: gap of unknown length
* * * 73715 75185: contig of 1471 bp in length
* * * 75186 75285: gap of unknown length
* * * 75286 76854: contig of 1569 bp in length
* * * 76855 76954: gap of unknown length
* * * 76955 78413: contig of 1459 bp in length
* * * 78414 78513: gap of unknown length
* * * 78514 79964: contig of 1451 bp in length
* * * 79965 80064: gap of unknown length
* * * 80065 81745: contig of 1861 bp in length
* * * 81746 81845: gap of unknown length
* * * 81846 83417: contig of 1572 bp in length
* * * 83418 83517: gap of unknown length
* * * 83518 84910: contig of 1393 bp in length
* * * 84911 85010: gap of unknown length
* * * 85011 86788: contig of 1778 bp in length
* * * 86789 86888: gap of unknown length
* * * 86889 88932: contig of 2004 bp in length
* * * 88933 88992: gap of unknown length
* * * 88993 90671: contig of 1678 bp in length
* * * 90672 90770: gap of unknown length
* * * 90771 93207: contig of 2437 bp in length
* * * 93208 93307: gap of unknown length
* * * 93308 96558: contig of 3351 bp in length
* * * 96559 96758: gap of unknown length
* * * 96759 100576: contig of 3818 bp in length
* * * 100577 104988: gap of unknown length
* * * 104989 105088: contig of 4312 bp in length
* * * 105089 108993: contig of 3805 bp in length
* * * 108994 108993: gap of unknown length
* * * 108994 112632: contig of 3639 bp in length
* * * 112633 112732: gap of unknown length
* * * 112733 116328: contig of 3596 bp in length
* * * 116329 116428: gap of unknown length
* * * 116429 120980: contig of 4552 bp in length
* * * 120981 121080: gap of unknown length
* * * 121081 126821: contig of 5741 bp in length
* * * 126822 126921: gap of unknown length
* * * 126922 131331: contig of 4410 bp in length
* * * 131332 131331: gap of unknown length
* * * 131332 134332: contig of 7274 bp in length
* * * 134332 138705: gap of unknown length
* * * 138706 138805: gap of unknown length
* * * 138806 145169: contig of 6364 bp in length
* * * 145170 145269: gap of unknown length
* * * 145270 150696: contig of 5427 bp in length
* * * 150697 150796: gap of unknown length
* * * 150797 159988: contig of 9192 bp in length
* * * 159989 160088: gap of unknown length
* * * 160089 167244: contig of 7156 bp in length
* * * 167245 167344: gap of unknown length
* * * 167345 175370: contig of 8026 bp in length.
* * * * * Location/Qualifiers
* * * * * 1..175370
* * * * * /organism="Pan troglodytes"
* * * * * /mol_type="genomic DNA"
* * * * * /db_xref="taxon:9598"
* * * * * /chromosome="UNK"
* * * * * /clone="RP43-21D9"
* * * * * 1..5345
* * * * * /note="assembly_name:Contig100"

misc_feature
source
FEATURES
1..175370
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="UNK"
/clone="RP43-21D9"
1..5345
/note="assembly_name:Contig100"

5446..17195
/note="assembly_name:Contig101"
17296..33250
/note="assembly_name:Contig102"
33351..45646
/note="assembly_name:Contig103"
45747..61992
/note="assembly_name:Contig104"
62093..63423
/note="assembly_name:Contig16"
63524..64666
/note="assembly_name:Contig3"
64767..65994
/note="assembly_name:Contig45"
66095..67430
/note="assembly_name:Contig1948"
67531..68982
/note="assembly_name:Contig5"
69083..70891
/note="assembly_name:Contig56"
70992..72172
/note="assembly_name:Contig960"
72273..73614
/note="assembly_name:Contig67"
73715..75185
/note="assembly_name:Contig68"
75286..76854
/note="assembly_name:Contig69"
76955..78413
/note="assembly_name:Contig73"
78514..79964
/note="assembly_name:Contig75"
80065..81745
/note="assembly_name:Contig77"
81846..83417
/note="assembly_name:Contig79"
83518..84910
/note="assembly_name:Contig980"
85011..86788
/note="assembly_name:Contig81"
86889..88992
/note="assembly_name:Contig82"
88993..90670
/note="assembly_name:Contig83"
90771..93207
/note="assembly_name:Contig84"
93308..96558
/note="assembly_name:Contig85"
96759..100576
/note="assembly_name:Contig86"
100677..104988
/note="assembly_name:Contig87"
105089..108993
/note="assembly_name:Contig88"
108994..112632
/note="assembly_name:Contig89"
112733..116328
/note="assembly_name:Contig90"
116429..120980
/note="assembly_name:Contig91"
121081..126821
/note="assembly_name:Contig92"
126922..131331
/note="assembly_name:Contig93"
131432..138705
/note="assembly_name:Contig94"
138806..145169
/note="assembly_name:Contig95"
145270..150696
/note="assembly_name:Contig96"
150797..159988
/note="assembly_name:Contig97"
160089..167244
```


misc_feature		/note="assembly_name:Contig98"	
167345..175370		/note="assembly_name:Contig99"	
ORIGIN			
Query Match	94.4%	Score 1599.8	DB 2; Length 175370;
Best Local Similarity	97.8%	Pred. No. 0;	
Matches 1668;	Conservative 0;	Mismatches 22;	Indels 15; Gaps 4;
QY	1	ATATACACTCTGTGATAA	CTACTCTGAAAGAGTGGGTCTGGAGACATAGATCCCAACAAG 60
Db	167982	ATATACACTCTGTGATAA	CTACTCTGAAAGAGTGGGTCTGGAGACATAGATCCCAACAAG 168041
QY	61	GAACCTGCTCCGGGATGAAA	CGTCATTTCAATAGGATCTTCCTGCCCAACATCTAC 120
Db	168042	GAACCTGCTCCGGGATGAAA	CGTCATTTCAATAGGATCTTCCTGCCCAACATCTAC 168101
QY	121	TTTCATCATCTTCTTGACT	TGGCATAGTCGGCAATGGATTGGTGATCCTGTGTCATGGGTAC 180
Db	168102	TTTCATCATCTTCTTGACT	TGGCATAGTCGGCAATGGATTGGTGATCCTGTGTCATGGGTAC 168161
QY	181	CAGAAGAAGCTAAGGAGCA	TACGCGACAAGTACCGGCTGCACCTGTCTCAGTGGCTGACCTC 240
Db	168162	CAGAAGAAGCTAAGGAGCA	TACGCGACAAGTACCGGCTGCACCTGTCTCAGTGGCTGACCTC 168221
QY	241	CTCTTTGTGTCATCACAC	TCTCCCTTCTGGGCAGTTGATGCCATGCTGACTGGTACTTTGGG 300
Db	168222	CTCTTTGTGTCATCACAC	TCTCCCTTCTGGGCAGTTGATGCCATGCTGACTGGTACTTTGGG 168281
QY	301	AAATTTTGTGTPAAGCGT	CTCATATCTCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
Db	168282	AAATTTTGTGTPAAGCGT	CTCATATCTCTACACTGTCAACCTCTACAGCAGCGTTCTC 168341
QY	361	ATCTGGGCTTCATCAGCT	CGGCTGACCTGCCATTTGCCATTTGTCACGCCCAACACAGTCAA 420
Db	168342	ATCTGGGCTTCATCAGCT	CGGCTGACCTGCCATTTGCCATTTGTCACGCCCAACACAGTCAA 168401
QY	421	AGGCCAAGGAACCTGCTG	CGTGAAGGCAGTCTATGTGGCGCTCTGGATCCAGCCCTC 480
Db	168402	AGGCCAAGGAACCTGCTG	CGTGAAGGCAGTCTATGTGGCGCTCTGGATCCAGCCCTC 168461
QY	481	CTCTGACTATACCTGACT	TTCATCTTTGCCAGCTCAGCCAGGGGACATCAGTCAGGGG 540
Db	168462	CTCTGACTATACCTGACT	TTCATCTTTGCCAGCTCAGCCAGGGGACATCAGTCAGGGG 168521
QY	541	GATGACAGGTACATCTGT	GACCGCTTTACCCGATAGCTGTGGATGTGGTGTTCAA 600
Db	168522	GATGACAGGTACATCTGT	GACCGCTTTACCCGATAGCTGTGGATGTGGTGTTCAA 168581
QY	601	TTCCAGCATATAATGGTGG	GTCTCATCTCTGCCGGCATCTGTCATCTCTCTGTCTACTGC 660
Db	168582	TTCCAGCATATAATGGTGG	GTCTCATCTCTGCCGGCATCTGTCATCTCTCTGTCTACTGC 168641
QY	661	ATCATCATCTTATAGCTGT	CACTTCCAGGGCCACAGAGCGCAAGGCCCTCAGAGC 720
Db	168642	ATCATCATCTTATAGCTGT	CACTTCCAGGGCCACAGAGCGCAAGGCCCTCAGAGC 168701
QY	721	ACAGTCATCTCCATCTAG	CTTTCTTTGCTGTGGCTGCCATATTTATGTGGGGATCAGC 780
Db	168702	ACAGTCATCTCCATCTAG	CTTTCTTTGCTGTGGCTGCCATATTTATGTGGGGATCAGC 168761
QY	781	ATCGACTCTTTCATCTTT	TGGGATGTCATCAGCAGGATGTGACTTCGAGAGCATTTGTG 840
Db	168762	ATCGACTCTTTCATCTTT	TGGGATGTCATCAGCAGGATGTGACTTCGAGAGCATTTGTG 168821
QY	841	CACAAGTGGATCTCCATC	CACAGAGGCCCTGGCTTCTTCCACTGTTGCTGAACCCCATC 900
Db	168822	CACAAGTGGATCTCCATC	CACAGAGGCCCTGGCTTCTTCCACTGTTGCTGAACCCCATC 168881
QY	901	CTCTATGCTTCTCTCGGG	CCAAAGTCTCTGCCAGCATGCACTCAACTCCATG 960
Db	168882	CTCTATGCTTCTCTCGGG	CCAAAGTCTCTGCCAGCATGCACTCAACTCCATG 168941

Parkway, St. Louis, MO 63108, USA

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M BA0480N14
----- Summary Statistics -----

Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly: Dye-terminator Big Dye; version 0.990319
Consensus quality: 134611 bases at least Q40
Consensus quality: 136848 bases at least Q30
Consensus quality: 138294 bases at least Q20
Insert size: 132000; agarose-fp
Insert size: 232319; sum-of-contigs
Quality coverage: 13.47 in Q20 bases; agarose-fp
Quality coverage: 9.68 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1119: contig of 1119 bp in length
* 1120: gap of unknown length
* 1220: contig of 1418 bp in length
* 2638: gap of unknown length
* 2738: contig of 2322 bp in length
* 5059: gap of unknown length
* 5159: gap of unknown length
* 6203: contig of 1044 bp in length
* 6303: gap of unknown length
* 6204: contig of 5511 bp in length
* 11814: gap of unknown length
* 11815: contig of 6539 bp in length
* 11915: gap of unknown length
* 18453: contig of 15131 bp in length
* 18553: gap of unknown length
* 33684: gap of unknown length
* 33685: contig of 23929 bp in length
* 33785: gap of unknown length
* 57713: gap of unknown length
* 57813: contig of 26738 bp in length
* 84551: gap of unknown length
* 84552: contig of 37634 bp in length
* 122286: gap of unknown length
* 122386: contig of 62830 bp in length.
Location/Qualifiers

1. 185215
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="R223-480N14"

1. 1119
/note="assembly_name:Contig51"
1220. 2637
/note="assembly_name:Contig103"
2738. 5059
/note="assembly_name:Contig122"
5160. 6203
/note="assembly_name:Contig124"
6304. 11814
/note="assembly_name:Contig126"
11915. 18453
/note="assembly_name:Contig127"
18554. 33684

FEATURES

Source

misc_feature
1. 1119
/note="assembly_name:Contig51"
1220. 2637
/note="assembly_name:Contig103"
2738. 5059
/note="assembly_name:Contig122"
5160. 6203
/note="assembly_name:Contig124"
6304. 11814
/note="assembly_name:Contig126"
11915. 18453
/note="assembly_name:Contig127"
18554. 33684

/note="assembly_name:Contig128"
33785. 57713
/note="assembly_name:Contig129"
57814. 84551
/note="assembly_name:Contig130"
84652. 122285
/note="assembly_name:Contig131"
122386. 185215
/note="assembly_name:Contig132"

ORIGIN

Query Match 94.4%; Score 1599.8; DB 2; Length 185215;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1668; Conservative 0; Mismatches 22; Indels 15; Gaps 4;

QY 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGCTCTGGAGACTATGATCTCCAAACAG 60
Db 91321 ATATACACTTCTGATACTACTCTGAAGAAGTGGGCTCTGGAGACTATGATCTCCAAACAG 91380

QY 61 GAACCCCTGCTCCGGGATGAAACGTCCTCAATTTCAATAGGATCTTCTCCGCCACCATCTAC 120
Db 91391 GAAACCCCTGCTCCGGGATGAAACGTCCTCAATTTCAATAGGATCTTCTCCGCCACCATCTAC 91440

QY 121 TTATCATCTTCTTACTGCTGGCATAGTGGCAATGGATGGTATGATCTTGGTATCGATGGGTAC 180
Db 91441 TTATCATCTTCTTACTGCTGGCATAGTGGCAATGGATGGTATGATCTTGGTATCGATGGGTAC 91500

QY 191 CAGAAGAGCTTAAGGAGCATGACGACAGTACCGGCTGCACCTGTGACGTGGCTGACCTC 240
Db 91501 CAGAAGAGCTTAAGGAGCATGACGACAGTACCGGCTGCACCTGTGACGTGGCTGACCTC 91560

QY 241 CTCTTTGTCATCACATCCCTTCTGGGAGTTGATGCCATGGCTGATCTGTGATCTTTGGG 300
Db 91561 CTCTTTGTCATCACATCCCTTCTGGGAGTTGATGCCATGGCTGATCTGTGATCTTTGGG 91620

QY 301 AAATTTTGTGTAAGGCTGTCCATATCATCTATCATCTGTCAACCTCTCAGCAGGGTTCTC 360
Db 91621 AAATTTTGTGTAAGGCTGTCCATATCATCTATCATCTGTCAACCTCTCAGCAGGGTTCTC 91680

QY 361 ATCTGGGCTTTCATCAGCTGGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCAA 420
Db 91681 ATCTGGGCTTTCATCAGCTGGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCAA 91740

QY 421 AGGCCAAGAAATCTGTGGCTGAAAGAGCTATGTGGGCGTCTGGATCCAGCCCTC 480
Db 91741 AGGCCAAGAAATCTGTGGCTGAAAGAGCTATGTGGGCGTCTGGATCCAGCCCTC 91800

QY 481 CTCCTGACTATACCTGACTTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGG 540
Db 91801 CTCCTGACTATACCTGACTTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGG 91860

QY 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTGTTCAA 600
Db 91861 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTGTTCAA 91920

QY 601 TTCCAGCATATAATGGTGGTCTCATCTCTGCCGCGCATCGTCATCTCTCTCTGTACTGC 660
Db 91921 TTCCAGCATATAATGGTGGTCTCTCTCTGCCGCGCATCGTCATCTCTCTCTGTACTGC 91980

QY 661 ATCATCATCTCTAAGCTGTCACTCCAAAGGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 720
Db 91981 ATCATCATCTCTAAGCTGTCACTCCAAAGGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 92040

QY 721 ACAGTCATCTCTATCTGTAGCTTTCTTGGCTGTGGTGGCCATATTATGTGGGGATCAGC 780
Db 92041 ACAGTCATCTCTATCTGTAGCTTTCTTGGCTGTGGTGGCCATATTATGTGGGGATCAGC 92100

QY 781 ATCGACTCTCTCATCTCTTTTGGGAGTCTATCAAGCAAGGATGTGATTTGAGAGCAATGTG 840
Db 92101 ATCGACTCTCTCATCTCTTTTGGGAGTCTATCAAGCAAGGATGTGATTTGAGAGCAATGTG 92160

QY 841 CACAAGTGGATCTCCATCAAGAGCCCTCTCTTCCACTGTGCTGTGACCCCATC 900
Db 841 CACAAGTGGATCTCCATCAAGAGCCCTCTCTTCCACTGTGCTGTGACCCCATC 900

[illegible]

Qy	541	GATGACAGGTACATCTGTGTGACCGCCCTTTACCCCGATAGCCTGTGGAGTGTGGTGTTCAA	600
Db	633	GATGACAGGTACATCTGTGTGACCGCCCTTTACCCCGATAGCCTGTGGAGTGTGGTGTTCAA	692
Qy	601	TTCCAGCATATAATGTGTGGGTCTCATCTCCCGGATCGTCACTCTCTCTGTTACTGC	660
Db	693	TTCCAGCATATAATGTGTGGGTCTCGTCTCTCCCGCATCGTCACTCTCTCTGTTACTGC	752
Qy	661	ATCATCATCTCTAAGCTGTCACTCCAAAGGCCACGAGAAGCGCAAGGCCCTCAAGACG	720
Db	753	ATCATCATCTCTAAGCTGTCACTCCAAAGGCCACGAGAAGCGCAAGGCCCTCAAGACG	812
Qy	721	ACAGTCATCTCATCTTAGCTTTCTTTTGCTGCTGCTGCTGCATATTATGTGGGGATCAGC	780
Db	813	ACAGTCATCTCATCTTAGCTTTCTTTTGCTGCTGCTGCTGCATATTATGTGGGGATCAGC	872
Qy	781	ATCGACTCCTTCATCTCTTTGGGAGTCATCAAGCAAGGATGTCACTTCGAGAGCATTTG	840
Db	873	ATCGACTCCTTCATCTCTTTGGGAGTCATCAAGCAAGGATGTCACTTCGAGAGCATCGTG	932
Qy	841	CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTCTTCCACTGTGGCTGAACCCGATC	900
Db	933	CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTCTTCCACTGTGGCTGAACCCGATC	992
Qy	901	CTCTATGCCCTTCCTCGGGGCAAGTTCAAAGGCTCTGCCAGCATGCACCTCAACTCCATG	960
Db	993	CTCTATGCCCTTCCTCGGGGCAAGTTCAAAGGCTCTGCCAGCATGCACCTCAACTCCATG	1052
Qy	961	AGCAGAGGCTCCAGCTCCAAAGATCTTTTCCAAAGGAAGCGGGTGGACACTCTTCGCTC	1020
Db	1053	AGCAGAGGCTCCAGCTCCAAAGATCTTTTCCAAAGGAAGCGGGTGGACACTCTTCGCTC	1112
Qy	1021	TCCACGGAGTCAGAACTCCTCAGTTTTTCACTCCAGCTTAACCCCTATTGTCAAAGACTTAT	1080
Db	1113	TCCACGGAGTCAGAACTCCTCAGTTTTTCACTCCAGCTTAACCTATTGTCAAAGACTATAT	1172
Qy	1081	A----ATATATATATATATATATAAGAACTTTTTTATGTATACATTTTCAGATATA	1136
Db	1173	AAATATATATATATATATATATAAGAACTTTTTTATGTATACATTTTCAGATATA	1232
Qy	1137	AGAGACTGACCGAGTCCTGTACAGTTTTTTTTTTTTTAATTGACTGTTCGGAGTTTATG	1196
Db	1233	AGAGACTGACCGAGTCCTGTACAGTTTTTTTTTTTTT-----ATTGACTGTTCGGAGTTTATG	1287
Qy	1197	TTCCCTAGTTTTTGTGAGGTTTGACCTAAATTTATATAAAATA-----TTGTTTTTGTGTTG	1252
Db	1288	TTCCCTAGTTTTTGTGAGGTTTGACCTAAATTTATATAAAATACCTTTTTTTTGTGTTGTTG	1347
Qy	1253	TTTCATGTGAATGAGCTGTAGGCAGACCTGTGGCCAGTTCTTAGTAGCTGTTTATCT	1312
Db	1348	TTTCATGTGAATGAGCTGTAGGCAGACCTGTGGCCAGTTCTTAGTAGCTGTTTATCT	1407
Qy	1313	GTGTGTAGGACTGTAGAACTGTAGCAAGAAACTGAACATTCAGAAATCTGTGTGTAAT	1372
Db	1408	GTGTGTAGGACTGTAGAACTGTAGCAAGAAACTGAACATTCAGAAATCTGTGTGTAAT	1467
Qy	1373	TGAATAAAGCTAGCGGTGATCCTCAGCTGTGTGTCGATAATCTCTTCATTCGAGGAGCA	1432
Db	1468	TGAATAAAGCTAGCGGTGATCCTCAGCTGTGTGTCGATAATCTCTTCATTCGAGGAGCA	1527
Qy	1433	CCCCACCCCAACCCCAACCCCAACCCCACTTAAATTTGTTGTTATGCTGTGTGATGG	1492
Db	1528	CCCCACCCCAACCCCAACCCCAACCCCACTTAAATTTGTTGTTATGCTGTGTGATGG	1587
Qy	1493	TTTGTGTTGGTTTTTTTTT---GTTGTGTTGTTTTTTTTTTTTTCTGTAAGAATGGCAC	1549
Db	1588	TTTGTGTTGGTTTTTTTTT---GTTGTGTTGTTTTTTTTTTTTTCTGTAAGAATGGCAC	1647
Qy	1550	TTAAACCAGAGCCTGAATGGTGGTAGAATGCTGGGTTTTTTTTTTGTTGTTGTTTTT	1609
Db	1648	TTAAACCAGAGCCTGAATGGTGGTAGAATGCTGGGTTTTTTTTTTGTTGTTGTTTTT	1707
Qy	1610	TTCAAGTGTAGTATGACTTCAGTCCCTCAAAATGTACAGCTCTGTATATACATT	1669

DB 1708 TTACAGTTTTCAGAGTAGATGATTTACCTCCCTACAATGTACAGTCTTGATTACATT 1767

QY 1670 GTTAATAAAAAAGTCAATGATAAAGTT 1694

DB 1768 GTTAATAAAAAAGTCAATGATAAAGTT 1792

RESULT 6

MMLESTRPT 1575 bp mRNA linear ROD 11-NOV-1996

LOCUS M.musculus mRNA for leukocyte-derived seven transmembrane domain

DEFINITION receptor.

ACCESSION X95582

VERSION X95582.1 GI:1666648

KEYWORDS lestr gene; leukocyte-derived seven transmembrane domain receptor.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Moepps, B., Prodi, R., Kessler, H. and Gierschik, P. cDNA cloning and genomic organization of a leukocyte-derived seven transmembrane domain receptor (LESTR) from mouse: a murine homologue of the human HIV-1 entry cofactor fusin

AUTHORS Unpublished

TITLE 2 (bases 1 to 1575)

JOURNAL Moepps, B.

REFERENCE Direct Submission

AUTHORS Submitted (19-JUL-1996) B. Moepps, Universitaet Ulm.

TITLE Pharmacology/Toxicology, Albert-Einstein Allee 11, D-89081 Ulm, FRG

JOURNAL Revised by author 11-NOV-96

REMARK On Nov 12, 1996 this sequence version replaced gi:1657351.

COMMENT Location/Qualifiers

FEATURES

source 1..1575

1..1575

/organism="Mus musculus"

/mol_type="mRNA"

/strain="B6CBA"

/db_xref="taxon:10090"

/haplotype="diploid"

/sex="female"

/tissue_type="thymus"

/dev_stage="6-8 weeks old"

gene 1..1575

1..1575

/gene="lestr"

CDS 90..1169

/gene="lestr"

/codon_start=1

/product="leukocyte-derived seven transmembrane domain receptor"

/protein_id="CAA67894.1"

/db_xref="GI:11666648"

/db_xref="GOA:P70658"

/db_xref="SWISS-PROT:P70658"

/translation="MEPISVISIYTSNSEEVSQGYDSNKEPCFCDENVFNRIFLP
TIVPIELTGIVGNGLVIMVQKLRMSMTDKYRLHLSVADLLFVITLFFWADVAMA
DWYFELCKAVHIIYVNLVSVLILAFISLDYLAIVHATNSQRPKRLLAEKAVY
GWTPALLITPDPIFADVSQDIDISQDDRYICDRLYPDSLMVWVFOFHIMVGLVLP
GIVILSCYCIILSKLSKHGQEKALKTKVILIAFFACWLPVYTGISIDSFILIG
IKQCFDESIVHKWISITELAFHFHCLNFIILAFGLKFKZSAQHAINLSKSGSSLK
ILSKRGKHSHSVSTSESSSFHSS"

polya_signal 1498..1503

/gene="lestr"

ORIGIN

Query Match 81.7%; Score 1383.8; DB 10; Length 1575;

Best Local Similarity 98.3%; Pred. No. 2.8e-233; Indels 13; Gaps 3;

Matches 1434; Conservative 0; Mismatches 12;

QY 1 ATATACATCTTGATAACTACTCTGAGAGAGTGGGGTCTCGAGACTATGACTCCAAAG 60

DB 111 ATATACATCTTGATAACTACTCTGAGAGAGTGGGGTCTCGAGACTATGACTCCAAAG 170

	/db_xref="GI:1666647"	
	/db_xref="GOA:P70658"	
	/db_xref="Swiss-Prot:P70658"	
	/translation="MEPISVIVTQKLSRMTDKYRLHLSVADLLFVLTPFWAVDAMA TIYFLITGTVGNGLVILVMGYTKLSDRYLAIVHAITNSORPKLIAEKAVYV DWYFKELCKRAVIIIVNLYSVLLAFSLDRYLAIHATNSORPKLIAEKAVYV GWIPALLIITPEIFPADVSQGISQGGDDRYICDLRPDSLMNVVPQPHIMVGLVLP GVILSYCYIIISKLSSHGKKRKALKUTIVILIAFPACWLFYIVGISIDSLILGV IKGCCDESIVHKWISITEALAPFHCCNLFIYAFLGAKPKSAQHIALMSMRSGSLK ILSKRGKGSHSVSTESSESPHSS"	
misc_feature	/gene="lestr" 105..2305	
intron	/note="splice variant/intron" 111..2305	
exon	/gene="lestr" /number=1 2306..>3770	
polyA_signal	/number=2 3693..3698	
ORIGIN		
Query Match	81.7%; Score 1383.8; DB 10; Length 3770;	
Best Local Similarity	98.3%; Pred.No. 3e-293;	
Matches 143; Conservative	0; Mismatches 12; Indels 13; Gaps 3;	
Qy	1 ATATACACTCTTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAACAAG 60	
Dd	2306 ATATACACTCTTGATAACTACTCTGAAGAAGTGGGTTCGGAGACTATGACTCCAACAAG 2365	
Qy	61 GAACCTCGCTCCGGGATGAAGAGTCCATTCAATAGGATCTTCTGCCCAACATCTTAC 120	
Dd	2366 GAACCTCGCTCTCCGGGATGAAGAGTCCATTCTTAATAGGATCTTCTGCCCAACATCTTAC 2425	
Qy	121 TTCAATCATCTTTTGAATGCGGCATAGTCGGCAATGGATTTGGTGTGATCCTGGTTCATGGGGTTAC 180	
Dd	2426 TTCAATCATCTTTTGAATGCGGCATAGTCGGCAATGGATTTGGTGTGATCCTGGTTCATGGGGTTAC 2485	
Qy	181 CAGNAGAGCTTAGGACATGACGGACACAGTACCGGCTGCACCTGTCACTGGCTGACCTC 240	
Dd	2486 CAGAAGAAGCTTAGGAGCATGACGGACAAAGTACCGGCTGCACCTGTCACTGGCTGACCTC 2545	
Qy	241 CTCCTTTGTCATCACACTCCCCTTCTGGGCGAGTTGATGCCATTTGCTGACTGGTACTTTGGG 300	
Dd	2546 CTCCTTTGTCATCACACTCCCCTTCTGGGCGAGTTGATGCCATTTGCTGACTGGTACTTTGGG 2605	
Qy	301 AAATTTTGTGAAGGTGTCCATATCATCTCACTGTCAAACCTTTACAGACGGTTCTC 360	
Dd	2606 AAATTTTGTGAAGGTGTCCATATCATCTCACTGTCAAACCTTTACAGACGGTTCTC 2665	
Qy	361 ATCTCGGCTTCATCAGCCTGGACCGGTACTCGCCATTGTCCACGCCAACCAACAGTCAA 420	
Dd	2666 ATCTCGGCTTCATCAGCCTGGACCGGTACTCGCTATTGTCCACGCCAACCAACAGTCA 2725	
Qy	421 AGGCCAAGGAAACTGTGGCTGAAAAAGCAGTCTATGTGGGGCTGTGGATCCAGGCCCTC 480	
Dd	2726 AGGCCAAGGAAACTGTGGCTGAAAAAGGCACTCTATGTGGGGCTGTGGATCCAGGCCCTC 2785	
Qy	481 CTCTGACTATACCTGACTTTTCATCTTTTGGCGAGTCTGACCATTGTCCACGCCGAGCATCACTCAGGGG 540	
Dd	2786 CTCTGACTATACCTGACTTTTCATCTTTTGGCGAGTCTGACCAGGCGAGCATCACTCAGGGG 2845	
Qy	541 GATGACAGGTACATCTGTGACCGCCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA 600	
Dd	2846 GATGACAGGTACATCTGTGACCGCCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA 2905	
Qy	601 TTCAGCATATATGGTGGGTCTCATCTCGCCGGGATCGTCACTCTCTCTCTTACTGC 660	
Dd	2906 TTCAGCATATATGGTGGGTCTCGTCTCTCGCCGGGATCGTCACTCTCTCTCTTACTGC 2965	
Qy	661 ATCATCATCTCTAAGCTGTCACTCCAAGGCCAACAGAGCGCCCTTCAGAAGCG 720	
Dd	2966 ATCATCATCTCTAAGCTGTCACTCCAAGGCCAACAGAGCGCCCTTCAGAAGCG 3025	

Cardenas, V., Carter, K., Cavazos, J., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Dexamo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louissegh, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mathiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelimeh, O., Okwunigbo, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlecyk, R., Wood, H., Worely, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 190149)
Worley, K.C.

Direct Submission
Submitted (22-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 190149)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23269619.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GXCR
Center clone name: CH230-108A24
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 179135 bases at least Q40
Consensus quality: 181484 bases at least Q30
Consensus quality: 183011 bases at least Q20
Estimated insert size: 187231; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 40067: contig of 40067 bp in length
* 40068 40167: gap of unknown length
* 114661 114760: contig of 74493 bp in length
* 114661 114760: gap of unknown length
* 114761 188186: contig of 73426 bp in length
* 188187 188287: gap of unknown length
* 188287 190149: contig of 1863 bp in length.
* Location/Qualifiers
* source
* 1..190149
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-108A24"

Query Match 69.0%; Score 1168.2; DB 2; Length 190149;
Best Local Similarity 84.6%; Pred. No. 9.1e-246;
Matches 1456; Conservative 0; Mismatches 143; Indels 122; Gaps 8;

QY 1 ATATACACTTCTGATAACTACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCCAACAG 60
DB 20662 ATATACACTTCTGATAACTACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCCAACAG 20721

QY 61 GAACCTGTCTCCGGATGAAACGTCCTTCATAGGATCTTCTGCCACCATCTAC 120
DB 20722 GAACCTGTCTCCGGATGAAACGTCCTTCATAGGATCTTCTGCCACCATCTAT 20781

QY 121 TTATCATCTTCTGACTGCGCATAGTGGGCAATGGTTCGATCTGCTGCTGCTTAC 180
DB 20782 TTATCATCTTCTGACTGCGCATAGTGGGCAATGGTTCGATCTGCTGCTGCTTAC 20841

QY 181 CAGAAGAGTAAAGGAGCATGACGGAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
DB 20842 CAGAAGAGTAAAGGAGCATGACGGAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 20901

QY 241 CTCTTTGTCTACATCTCCCTTCTCGGAGTGTGATGCGCATGGCTGCTGCTGCTTGGG 300
DB 20902 CTCTTTGTCTACATCTCCCTTCTCGGAGTGTGATGCGCATGGCTGCTGCTGCTTGGG 20961

QY 301 AAATTTTGTGTAAAGGTGTCCATATCATCTACACTGTCAACCTCTACAGACGCTTTC 360
DB 20962 AAATTTTGTGTAAAGGTGTCCATATCATCTACACTGTCAACCTCTACAGACGCTTTC 21021

QY 361 ATCTGCGCTTCTACAGCTGCGGAGTGTGATGCGCATGGCTGCTGCTGCTTGGG 420
DB 21022 ATCTGCGCTTCTACAGCTGCGGAGTGTGATGCGCATGGCTGCTGCTGCTTGGG 21081

QY 421 AGGCCAAGGAAACCTGCTGGCTGAAAGAGGAGTCTATGTGGGCGTCTGCTGCTGCTT 480
DB 21082 AGGCCAAGGAGTGTGCTGGCTGAAAGAGGAGTCTATGTGGGCGTCTGCTGCTGCTT 21141

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

QY	481	CTCCTGACTATACCTTGACTTTCATCTTTGCCGACGTCAGCCAGGGGGACATCAGTCAGGGG	540
Db	21142	CTCCTGACTATACCTTGACTATCTTCGCCGATGTCAGCCAGGGGGAC-----	21189
QY	541	GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCTGTGGATGGTGGTGTTC	600
Db	21190	---GGCAGGTACATCTGTGACCGCCTTTACCCCGACAGCCTGTGGATGGTGGTGTTC	21246
QY	601	TTCCAGCATATAATGGTGGGTCTCATCTGTCCCGGCATCTGTCATCTCTCTGTACTGTC	660
Db	21247	TTCCAGCATATAATGGTGGGTCTCATCTGTCCCGGCATCTGTCATCTCTCTGTACTGTC	21306
QY	661	ATCATCATCTTAAGCTGTCACTTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG	720
Db	21307	ATCATCATCTTAAGCTGTCACTTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACT	21366
QY	721	ACAGTCATCCTCATCTAGCTTTCTTTGCTGCTGGCTGGCTGCCATATATGTGGGATCAGC	780
Db	21367	ACGGTTCATCCTTATCTGTGCTTTCTTTGCTGCTGGCTACCGTATATAGCTGGGATCAGC	21426
QY	781	ATCGACTCCTTCATCCTTTTGGGAGTCTCAAGCAAGGATGTGACTTCGAGAGCATTTGTG	840
Db	21427	ATCGATTCCTTCATCCTTTTGGAGTCTCAAGCAAGGATGTGACTTCGAGAGCATTTGTG	21486
QY	841	CACAAGTGGATCTCCATCA CAGAGGCCCTCGCTCTTTCCACTGTGTGCTGAAACCCCATC	900
Db	21487	CACAAGTGGATCTCCATCA CAGAGGCCCTCGCTCTTTCCACTGTGTGCTGAAACCCCATC	21546
QY	901	CTGTATGCTTCTCTGGGGCCAAAGTTCAAAGCTCTGCCAGCATGCACTCAATCCCATG	960
Db	21547	CTGTATGCTTCTCTGGGGCCAAAGTTCAAAGCTCTGCCAGCATGCACTCAATCCCATG	21606
QY	961	AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAACGGGGTGGACACTCTTTCGGTC	1020
Db	21607	AGCAGAGGCTCCAGCCTCAAGATCCTTTCCAAAGGAAACGGGGTGGACACTCTTTCGGTC	21666
QY	1021	TCCACGAGTTCAGAACTCTCCAGTTTTCACTCCAGCTAACCCCTTATGCAAGACTTATAT	1080
Db	21667	TCCACGAGTTCAGAACTCTCCAGTTTTCACTCCAGCTAACCACTTACGCAAGACATATAT	21726
QY	1081	AATATATATATATATAT-----GATAAAGAACTTTTTATCTTACACATTTTCCAG	1131
Db	21727	AATATATATATATATATATATATGACAAAGAACTTTTTATGTTACATTTCCAG	21786
QY	1132	ATATAAGAGACTGACAGTCTTTGACAGTTTTTTTTTTTTTAACTGTTGGAGT	1191
Db	21787	ATATAAAGAGACTGACCAATCTTTGTACAGTTTTTTTTTATGACTGTTGGAGTTATGTTTC	21846
QY	1192	TATGTTTCTCT-----AGTTTTGTGAGTTTTGACTTAATTTATATAATAT	1239
Db	21847	TCTAGTTTTGT	21906
QY	1240	-----TGTTTTTTTGTGTTTTCATGTGAATGAGCGTCTAGGACGAGCCTGTGGCCAGT	1293
Db	21907	TTTTTCTGTGTTTGTCTGTTTTCAGTGAATGAGTGTCTAGGACGAGCCTGTGGCCAGT	21966
QY	1294	TCTTAGTAGTGTTTATCTGT	1353
Db	21967	TCTTAGT	22026
QY	1354	TCCAGAAATGTGTGTTAAATGAATGAAGCTAGCCGTGATCTCCAGCTGTGTGTGATAAT	1413
Db	22027	TCCAGAAATGTGTGTTAAATTCGGATAAAGCTAGAGTACCTCAGCTGTGTGTGTGTGTGT	22086
QY	1414	CTCTTCATTCGAGGAGACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAATGTT	1473
Db	22087	CTCTT-----GGCCCGGCCCAACCCCGTTCCTGAATGTT	22120
QY	1474	TGTTATGCTGT	1533
Db	22121	TGTTGTGTGTGTGT-----GTTTCT	22141

Db 680 AGCCCAAGGAACTGCTGGCTGAAAGGAGTCTATGTGGGGCTCTGGATCCAGCCCTC 739
Qy 481 CTCCTGACTATACCTGACTTTCATCTTTTGGCGAGCTCAGCCAGGGGACATCATGTCAGGG 540
Db 740 CTCCTGACTATACCTGACTTTCATCTTTTGGCGAGCTCAGCCAGGGGACATCATGTCAGGG 799
Qy 541 GATGACAGGTACATCTGTGTGACCGCCCTTACCCCGATAGCCTGTGAGTGTGTGTTCAA 600
Db 800 GATGACAGGTACATCTGTGTGACCGCCCTTACCCCGATAGCCTGTGAGTGTGTGTTCAA 859
Qy 601 TTCACGATATATGTTGGTCTCATCTGCGCGGATGCTCATCTCTCTCTCTCTCTCTCTCTCT 660
Db 860 TTCACGATATATGTTGGTCTCATCTGCGCGGATGCTCATCTCTCTCTCTCTCTCTCTCTCT 919
Qy 661 ATCATCATCTCTAAGCTGTGCACACTCAAGGGGCCACGAGAGCGCAAGGCCCTCAAGAG 720
Db 920 ATCATCATCTCTAAGCTGTGCACACTCAAGGGGCCACGAGAGCGCAAGGCCCTCAAGAG 979
Qy 721 ACAGTCATCTCATCTAGCTTCTTTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 980 ACAGTCATCTCATCTAGCTTCTTTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039
Qy 781 ATCGACTCTCTTCTATCTCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGT 840
Db 1040 ATCGACTCTCTTCTATCTCTTTTGGGAGTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGT 1099
Qy 841 CACAAGTGTATCTCATACAGAGGCGCTGCGCTTCTTCCATGTTGCTGCTGCTGCTGCTGCTGCT 900
Db 1100 CACAAGTGTATCTCATACAGAGGCGCTGCGCTTCTTCCATGTTGCTGCTGCTGCTGCTGCTGCT 1159
Qy 901 CTCTATGCTCTCTCGGGGCAAGTTCAAAAGCTCTGCCAGATGCTGCCAGATGCTCAACTCCATG 960
Db 1160 CTCTATGCTCTCTCGGGGCAAGTTCAAAAGCTCTGCCAGATGCTGCCAGATGCTCAACTCCATG 1219
Qy 961 AGCAGAGGCTCCAGCTCAAGATCTCTTTTCCAAAGGAAAGCGGGGTGAGCACTCTCCGCTC 1020
Db 1220 AGCAGAGGCTCCAGCTCAAGATCTCTTTTCCAAAGGAAAGCGGGGTGAGCACTCTCCGCTC 1279
Qy 1021 TCCAGGAGTCAGATCTCCAGCTTCTCTCCAGATGCTTCCAGATGCTTCCAGATGCTTCCAGAT 1080
Db 1280 TCCAGGAGTCAGATCTCCAGCTTCTCTCCAGATGCTTCCAGATGCTTCCAGATGCTTCCAGAT 1339
Qy 1081 AATATATATATATATATGATAAAGAACTTTTATGTGTATACATATTTTCCAGATATTAAGAG 1140
Db 1340 AATATATATATATATATGATAAAGAACTTTTATGTGTATACATATTTTCCAGATATTAAGAG 1399
Qy 1141 ACTGACAGCTCTGTACAGTTT 1163
Db 1400 ACTGACAGCTCTGTACAGTTT 1422

RESULT 11
MMLCR12 1180 bp mRNA linear ROD 18-NOV-1997
LOCUS Mus musculus lcr-1 gene.
DEFINITION Z80112
ACCESSION 280112.1 GI:2632100
VERSION CXC chemokine receptor 4; CXCR-4; lcr-1 gene.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Schubel, A., Burgstahler, R. and Lipp, M.
The mouse homologue of the human HIV coreceptor CXCR-4 (FUSIN):
High expression in thymus and lymphoid tissues
Unpublished
JOURNAL 2 (bases 1 to 1180)
REFERENCE Lipp, M.
AUTHORS Direct Submission
TITLE Submitted (11-SEP-1996) Martin Lipp, Max-Delbrueck-Center for
JOURNAL Molecular Medicine, MDC, Robert-Rössle-Strasse 10, BERLIN-BUCH,
D-13122, GERMANY

COMMENT On Nov 21, 1997 this sequence version replaced gi:1542888.
FEATURES
source 1. 1180
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129SV"
/db_xref="taxon:10090"
/tissue_type="thymus"
/dev_stage="adult"
gene 1. 1180
/gene="lcr-1"
CDS 1..1074
/gene="lcr-1"
/function="receptor for chemokine SDF-1"
/standard_name="CXC chemokine receptor 4"
/codon_start=1
/product="CXCR-4"
/protein_id="CA02202.1"
/db_xref="GI:1542888"
/db_xref="GOA:P70658"
/translation="MEPISYTDNYSVEYSGDYDSNKBPCFRDENVHFNRLPFTI
YFIFLFGIVGNGLIVMGVQKRLSMTDKRLHLSVADLLFVITLPPFAVDAMADW
YFGFLCKAVHIIYTNLYSVLLILAFISLDRLVLAIVHATNSORPKRLAEKAVYGV
WIPALATIPDFIADYSQSDISQDDRYICDRLYPSLAWVVFQFOHIMVGLILPGI
VILSCYCIILSKLSHSGHQRKALKTTVILLILAFACWLPYVYVIGISIDFILLGVK
QKDFESIVHKWISITEALAFPHCCLLNPILYAFILGAFKRSQAHLNMSRGSILIL
SKRGKGHSVSTESSESSPHSS"
exon <1..15
/gene="lcr-1"
/number=1
exon 16..>1180
/gene="lcr-1"
/number=2
ORIGIN
Query Match 67.3%; Score 1139.8; DB 10; Length 1180;
Best Local Similarity 99.2%; Pred. No. 1.1e-239;
Matches 1156; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
Qy 1 ATATACACTTCTGATTAAGTGGGTCTGAGAGTGGGTCTGAGACTATGATCTCAACAAG 60
Db 16 ATATACACTTCTGATTAAGTGGGTCTGAGAGTGGGTCTGAGACTATGATCTCAACAAG 75
Qy 61 GAACCTCTCTTCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGCCACCATCTAC 120
Db 76 GAACCTCTCTTCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGCCACCATCTAC 135
Qy 121 TTGATCATCTTCTTGAATGCGATAGTCGGCAATGATGATGATGATGATGATGATGATGATGAT 180
Db 136 TTGATCATCTTCTTGAATGCGATAGTCGGCAATGATGATGATGATGATGATGATGATGATGAT 195
Qy 181 CAGAGAGCTAAAGAGCATGACGCAAGTACCGGCTGACCTGTGAGTGGCTGAGCTGCCTC 240
Db 196 CAGAGAGCTAAAGAGCATGACGCAAGTACCGGCTGACCTGTGAGTGGCTGAGCTGCCTC 255
Qy 241 CTCCTTGTGATCAGACTCCCTTCTTGGGCACTTGTGATGATGATGATGATGATGATGATGATGAT 300
Db 256 CTCCTTGTGATCAGACTCCCTTCTTGGGCACTTGTGATGATGATGATGATGATGATGATGATGAT 315
Qy 301 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 316 AAATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 375
Qy 361 ATCTGGCTCTTCTGAGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 376 ATCTGGCTCTTCTGAGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
Qy 421 AGGCCAAGAAACTGCTGGTGAAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 436 AGGCCAAGAAACTGCTGGTGAAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
Qy 481 CTCCTGATATACCTGACTTTCATCTTTGCCGACGTCAGCCAGGGGACATCAGTCAGGGG 540

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.G., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Giang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshubew, L., Louisseghe, H., Lozador, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathewine, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitadas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekemeh, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Ffannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzio, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shwartsbeyn, A., Sisson, I., Sitter, C. D., Smajis, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wietczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 240931)
Workley,K.C.
Direct Submission
Submitted (16-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 240931)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23115278.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>) Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

REFERENCE
AUTHORS
TITLE
JOURNAL

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project information
Center project name: GKEX
Center clone name: CH230-175010

COMMENT

----- Summary Statistics			
Assembly program: Atlas 3.0;			
Consensus quality: 216793 bases at least Q40			
Consensus quality: 219254 bases at least Q30			
Consensus quality: 221224 bases at least Q20			
Estimated insert size: 228277; sum-of-contigs estimation			
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation			

* NOTE: Estimated insert size may differ from sequence length			
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).			
* NOTE: This is a 'working draft' sequence. It currently			
* consists of 1 contigs. Gaps between the contigs			
* are represented as runs of N. The order of the pieces			
* is believed to be correct as given, however the sizes			
* of the gaps between them are based on estimates that have			
* provided by the submitter.			
* This sequence will be replaced			
* by the finished sequence as soon as it is available and			
* the accession number will be preserved.			
* 1 240931: contig of 240931 bp in length.			
Location/Qualifiers			
1. .240931			
/organism="Rattus norvegicus"			
/mol_type="genomic DNA"			
/db_xref="taxon:10116"			
/clone="CH230-175010"			
1. .1249			
/note="wgs end extension			
clone_end:Sp6"			
complement(5161..5587)			
/note="clone boundary			
clone_end:Sp6			
site:EcoRI			
end sequence:BH357332"			
complement(232579..233220)			
/note="clone boundary			
clone_end:T7			
1. .237558			
site:EcoRI			
end sequence:BH357330"			
235468..237558			
/note="wgs end extension			
clone_end:T7"			
239450..240931			
/note="wgs end extension			
clone_end:T7"			
ORIGIN			
Query Match 65.8%; Score 1114; DB 2; Length 240931;			
Best Local Similarity 84.2%; Pred. No. 7.4e-234; Indels 122; Gaps 8;			
Matches 1400; Conservative 0; Mismatches 140; Indels 122; Gaps 8;			
Qy	60	GGAAACCTGCTTCGGGATGAAACGTCATTTCAATAGGATCTTCTGCCCAACCACTTA	119
Db	237558	GGAAACCTGCTTCGGGATGAAACGTCATTTCAACAGGATCTTCTGCCCAACCACTTA	237499
Qy	120	CTTATCATCTCTTGACTGCGATGCGCAATGGATTGGTATCTGCTGGTCAATGGTTA	179
Db	237498	TTTATCATCTCTTGACTGCGATGCGCAATGGATTGGTATCTGCTGGTCAATGGTTA	237439
Qy	180	CCAGAAAGCTAAGGAGCATGACGCAAGTACCGCTGACCTGTCAGTGGCTGACCT	239
Db	237438	CCAGAAAGCTAAGGAGCATGACGCAAGTACCGCTGACCTGTCAGTGGCTGACCT	237379
Qy	240	CCTCTTGTCTATCACTCCCTTCTGGGCAATGATGCGCATGGCTGACTGTTTGG	299
Db	237378	CCTCTTGTCTATCACTCCCTTCTGGGCAATGATGCGCATGGCTGACTGTTTGG	237319
Qy	300	GAATTTTGTGTAAGGCTGCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCT	359
Db	237318	GAATTTTGTGTAAGGCTGCCATATCATCTACACCGTCAACCTCTACAGCAGCGTTCT	237259
Qy	360	CATCTGCGCTTCATCAGCTGGACCGGTACCTGCCATTGTCCACGCCACCAACAGTCA	419

Db	237258	CATCTGCGCTTCATCAGCCTGGACCGCTACCTTGGCCATTGTCCACGCCACCAACGCCA	237199
Qy	420	AAGGCCAAGGAACTGCTGCTGAAAGGCGAGTCTATGTGGGCTCTGGATCCCGCCT	479
Db	237198	GAGCCGAGGAACTGCTGCTGAAAGGCGCGTCTATGTGGGCTCTGGATCCCGCCT	237139
Qy	480	CCTCTGACTATACCTGACTTTCATCTTTGCGAGCTCAGCCAGGGGACATCATCAGGG	539
Db	237138	CCTCTGACTATACCTGACTTTCATCTTTGCGAGTCTCAGCCAGGGGAC	237090
Qy	540	GGATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGTTC	599
Db	237089	---GGCAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGTTC	237034
Qy	600	ATTCAGCATATAATGTTGGGTCTCATCTGCCCGGATCGTCACTCTCTCTCTCTCTG	659
Db	237033	GTTCCAGCAATCATGTGGGTCTCATCTGCCCGGATCGTCACTCTCTCTCTCTG	236974
Qy	660	CATCATCATCTCTAAGCTGTCACTCTCAAGGGCCACCAAGAGCGCAAGGCCCTCAAGAC	719
Db	236973	CATCATCATCTCTAAGCTGTCACTCTCAAGGGCCACCAAGAGCGCAAGGCCCTCAAGAC	236914
Qy	720	GACAGTCACTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	779
Db	236913	TACGGTCACTCTCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	236854
Qy	780	CATCGACTCTCTTCATCTCTTTGGGAGTCACTCAAGCAAGGATGTGACTTCGAGAGCAT	839
Db	236853	CATCGACTCTCTTCATCTCTTTGGGAGTCACTCAAGCAAGGATGTGACTTCGAGAGCAT	236794
Qy	840	GCACAGTGGATCTCCATCAGAGGCCCTGCGCTTCTTCCACTGTTGCCGTGAACCCCAT	899
Db	236793	GCACAGTGGATCTCCATCAGAGGCCCTGCGCTTCTTCCACTGTTGCCGTGAACCCCAT	236734
Qy	900	CCTCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCACTCAATCCAT	959
Db	236733	CCTCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCACTCAATCCAT	236674
Qy	960	GAGCAGAGCTCCAGCTCAAGATCTCTTCAAGAGAAAGGGGTGACACTCTTCCGT	1019
Db	236673	GAGCAGAGCTCCAGCTCAAGATCTCTTCAAGAGAAAGGGGTGACACTCTTCCGT	236614
Qy	1020	CTCCAGGAGTCAAGATCTCTCCAGTTCCTCAAGCTTATCCAGCTTACACATACGAA	1079
Db	236613	CTCCAGGAGTCAAGATCTCTCCAGTTCCTCAAGCTTATCCAGCTTACACATACGAA	236554
Qy	1080	TAAT	1130
Db	236553	TAAT	236494
Qy	1131	GATATAGAGCTGACCGCTCTTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	1190
Db	236493	GATATAGAGCTGACCGCTCTTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	236434
Qy	1191	TTTATGTTTCT	1239
Db	236433	CTCTAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	236374
Qy	1240	-----TGTTTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1292
Db	236373	TTTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	236314
Qy	1293	TTCTTAGTAGCTGTTTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1352
Db	236313	TTCTTAGTAGCTGTTTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	236254
Qy	1353	TTCCAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1412
Db	236253	CTCCAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	236194
Qy	1413	TTCTTCTCATTCGAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCCAC	1472
Db	236193	TTCTTCTCATTCGAGGAGCACCCACCCACCCACCCACCCACCCACCCACCCACCCAC	236160

/gene="CXCR-4"
 /number=1
 2308..3366
 /gene="CXCR-4"
 /number=2

exon
 ORIGIN

Query Match 62.2%; Score 1054.2; DB 10; Length 3366;
 Best Local Similarity 99.7%; Pred. No. 7.8e-231;
 Matches 1056; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	ATATACATCTTGTGATAACTACTCTGAAGAACTGGGGTCTGGAGACTATGACTCCAAACAAG	60
DB	2308	ATATACATCTTGTGATAACTACTCTGAAGAACTGGGGTCTGGAGACTATGACTCCAAACAAG	2367
QY	61	GAACCCCTGCTCCGGGATGAAAAAGTCCATTTCAATAGGATCTTCTGCCACCATCTAC	120
DB	2368	GAACCCCTGCTCCGGGATGAAAAAGTCCATTTCAATAGGATCTTCTGCCACCATCTAC	2427
QY	121	TTTCATCATCTTCTGACAGGCGATAGTCGGCAATGGATTGGTGATCCTCGTCATGGGTTAC	180
DB	2428	TTTCATCATCTTCTGACAGGCGATAGTCGGCAATGGATTGGTGATCCTCGTCATGGGTTAC	2487
QY	181	CAGAAAGAGCTAAGGAGCATGACGGCAAGTACCGGCTGCACCTGTCACTGGCTGACCTC	240
DB	2488	CAGAAAGAGCTAAGGAGCATGACGGCAAGTACCGGCTGCACCTGTCACTGGCTGACCTC	2547
QY	241	CCTTTGTGCATCACACTCCCTTCTGGGCAGTGTGATGCCATGGCTGACTGGTACTTTGGG	300
DB	2548	CCTTTGTGCATCACACTCCCTTCTGGGCAGTGTGATGCCATGGCTGACTGGTACTTTGGG	2607
QY	301	AAATTTTGTGTAAGGCTGTCCATATCATCTACCTGTCAACCTCTACAGACGGTTCTC	360
DB	2608	AAATTTTGTGTAAGGCTGTCCATATCATCTACCTGTCAACCTCTACAGACGGTTCTC	2667
QY	361	ATCCTGGCCTTCATCAGCGCTGACCGGTACCTCGGCATTGTCCAGCCACCAACAGTCAA	420
DB	2668	ATCCTGGCCTTCATCAGCGCTGACCGGTACCTCGGCATTGTCCAGCCACCAACAGTCAA	2727
QY	421	AGGCCAAGGAAACTGCTGGCTGAAAGGAGTGTATGTGGGGGTCTGGATCCAGCCCTC	480
DB	2728	AGGCCAAGGAAACTGCTGGCTGAAAGGAGTGTATGTGGGGGTCTGGATCCAGCCCTC	2787
QY	481	CTCCCTGACTATACCTGACTTTCATCTTTGCCAGCTGACGAGGGGGACATCAGTCAGGGG	540
DB	2788	CTCCCTGACTATACCTGACTTTCATCTTTGCCAGCTGACGAGGGGGACATCAGTCAGGGG	2847
QY	541	GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCGTGTGATGGTGGTGTTCAA	600
DB	2848	GATGACAGGTACATCTGTGACCGCCTTTACCCCGATAGCCGTGTGATGGTGGTGTTCAA	2907
QY	601	TTCCAGCATATTAATGGTGGGTCTCATCTGCCCGCATCGTCATCCTCTCCGTGTACTGC	660
DB	2908	TTCCAGCATATTAATGGTGGGTCTCATCTGCCCGCATCGTCATCCTCTCCGTGTACTGC	2967
QY	661	ATCATCATCTCTAAGGTGTACACTCCAGGGGCCACGAAAGCGCAAGGCCCTCAAGACG	720
DB	2968	ATCATCATCTCTAAGGTGTACACTCCAGGGGCCACGAAAGCGCAAGGCCCTCAAGACG	3027
QY	721	ACAGTCATCTCATCTAGCTTTCTTTGCCTGCTGGCTGCCATATTANGTGGGGATCAGC	780
DB	3028	ACAGTCATCTCATCTAGCTTTCTTTGCCTGCTGGCTGCCATATTANGTGGGGATCAGC	3087
QY	781	ATCGACTCTCCTTTTGGGGAGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGG	840
DB	3088	ATCGACTCTCCTTTTGGGGAGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGG	3147
QY	841	CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTGTGCTGAACCCCATC	900
DB	3148	CACAAGTGGATCTCCATCACAGAGGCCCTCGCCTTCTTCCACTGTGTGCTGAACCCCATC	3207
QY	901	CTCTATGCCCTTCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATG	960

3208	CTCTATGCTCTTCCTCGGGCGCAAGATTCAAAGAGCTGTCGCCAGCATGCATCAACTCCATG	3267
961	AGCAGAGGCTCCAGCCTCAAGATCCTTTTCAAAGGAAAGCGGGTGGACACACTTCCGTC	1020
3268	AGCAGAGGCTCCAGCCTCAAGATCCTTTTCAAAGGAAAGCGGGTGGACACACTTCCGTC	3327
1021	TCACCGAGTCAAGTCCTCCAGTCTTTTCACTCCAGCTAA	1059
3328	TCACCGAGTCAAGTCTCCAGTCTTTTCACTCCAGCTAA	3366
RESULT 14		
RN54791	1362 bp mRNA linear ROD 14-SEP-2001	
DEFINITION	Rattus norvegicus chemokine receptor LCR1 mRNA, complete cds.	
ACCESSION	U54791	
VERSION	U54791.1 GI:1354504	
KEYWORDS		
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
	Rattus.	
REFERENCE	1 (bases 1 to 1362)	
AUTHORS	Wong,M.L., Xin,W.W. and Duman,R.S.	
TITLE	Rat LCR1: cloning and cellular distribution of a putative chemokine receptor in brain	
JOURNAL	Mol Psychiatry 1 (2), 133-140 (1996)	
MEDLINE	97233286	
PUBMED	9118323	
REFERENCE	2 (bases 1 to 1362)	
AUTHORS	Wong,M.	
TITLE	Direct Submission	
JOURNAL	Submitted (11-APR-1996) Ma-Li Wong, Clinical Neuroendocrinology Branch, National Institutes of Health, NIMH, Building 10 Room 35231, 10 Center DR MSC 1284, Bethesda, MD 20892-1284, USA	
FEATURES	Location/Qualifiers	
source	1..1362	
	/organism="Rattus norvegicus"	
	/mol_type="mRNA"	
	/strain="Sprague-Dawley"	
	/db_xref="taxon:10116"	
	/clone="RLCR1"	
	44..1042	
	/codon_start=1	
	/product="chemokine receptor LCR1"	
	/protein_id="AAB01981.1"	
	/db_xref="GI:1354505"	
	/translation="MEIVTSDNYSEVSGDYDSNKEPCFRDENENFNRIPLTIYFI	
	IFELTGIVGNGLVILWGYOKLRSMWDKYRLHLISVADLLFVTLFPSDMAWYRGKPF	
	LCKAHVITVNLISVLLIATISLDRIYLAIVHATNSQSKALAEKADIIIFADVSQG	
	DERYCDRLPYSLMMVWFQFHWGLIPGVILISCYCIIISKLSHKGGKRAL	
	KTTVILIAFFACMLPYVVGISIDSPILLEVIKQGFESVHVHMIITEALAFPHCC	
	LNPLIYAFLGAKFKTSQAQHALNMSRGSLSLKILSKRGHGSVSTESSSSFHS"	
CDS		
ORIGIN		
Query Match	55.9%; Score 946.4; DB 10; Length 1362;	
Best Local Similarity	84.8%; Pred. No. 3.6e-197;	
Matches 1169; Conservative	0; Mismatches 116; Indels 93; Gaps 6;	
QY	1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCACAAG	60
DB	50 ATATACACTTCGGATTAAGTCTCGAAGAAGTAGGGTCTGGAGACTATGACTCCACAAG	109
QY	61 GAACCTTGCTTCGGGATGAAACGTCATTTCAATPAGGATCTTCTGCCCACCATCTAC	120
DB	110 GAACCTTGCTTCGGGATGAAACGCAAACTTCAACAGGATCTTCTGCCCACCATCTAT	169
QY	121 TTTCATCATCTTCTTGACTGTCATAGTCGCAATGGATTTGGTGATCTCTGTCATGGGTAC	180
DB	170 TTTATCATCTTCTTGACTGTCATAGTGGGCAATGGGTGTGTAATCTCTGTCATGGGTAC	229
QY	181 CAGAAGAAGCTAAGGAGCATGACGGCAAGATACCGCTGCACCTGTCAGTGGCTGACCTC	240

230	DB	CAGAGAAGCTGAGGAGCATGACAGACAAGTACCGGCTGCACCTGTCGGTGGCTGACCTC	289
241	QY	CTCTTTGTGTCATCACACTCCCTTCTTGGCGAGTTGATGCGATGGCTGACTGGTACTTTGGG	300
290	DB	CTCTTTGTGTCATCACACTCCCTTCTTGGCGAGTTGATGCGATGGCTGACTGGTACTTTGGG	343
301	QY	AAATTTTGTGTAAGGCTGTCCATATCATCTACATCTGTCAACTCTTACAGCAGCGTTCTC	360
344	DB	AAATTTTGTGTAAGGCTGTCCATATCATCTACATCTGTCAACTCTTACAGCAGCGTTCTC	403
361	QY	ATCTGGCCCTTCATCAGCCTGGACCGGTACCTTCGCCATTGTCCACGCCACCAACAGTCAA	420
404	DB	ATCTGGCCCTTCATCAGCCTGGACCGGTACCTTCGCCATTGTCCACGCCACCAACAGTCAA	463
421	QY	AGGCCAAGAAACTGCTGGCTGAAAGGCAAGTCTATGTGGGCGTCTGGATCCAGGCCCTC	480
464	DB	AGCGCAGGAAGCTGCTGGCTGAAAGG	491
481	QY	CTCCTGACTATACCTGACTTCATCTTTGGCCGAGCTCAGCCAGGGGACATCAGTCAGGGG	540
492	DB	-----CCGACATCATCTTCGCCGATGTCAGCCAGGGGAC-----	526
541	QY	GATGACAGTACATCTGTGACCGCCTTTATCCCGATAGCCTGTGGATGTGTTGTTCAA	600
527	DB	--GGCAGGTACATCTGTGACCGCCTTTATCCCGCAGCAGCCTGTGATGGTGTGTTCCAG	583
601	QY	TTCCAGCATATATGTTGGGTCTCATCTGCGCGGATCGTCACTCTCTCTGTTACTGTC	660
584	DB	TTCCAGCATATATGTTGGGTCTCATCTGCGCGGATCGTCACTCTCTCTGTTACTGTC	643
661	QY	ATCATCATCTCTAAGCTGTACATCTCAAGGGCCACCAGAAGCGCAAGCCCTCAAGACG	720
644	DB	ATCATCATCTCAAGCTGTACATCTCAAGGGCCACCAGAAGCGCAAGCCCTCAAGACT	703
721	QY	ACAGTCATCTCATCTAGCTTCTTTGGCTGCTGCTGCCATATATATGTTGGGGATCAGC	780
704	DB	ACGGTCATCTTATCTGCTCTTTTGGCTGCTGCTACCGTATTCAGTGGGGATCAGC	763
781	QY	ATCGACTCTCTCATCTCTTTGGGAGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG	840
764	DB	ATCGATTCTCTCATCTCTTTGGAGGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG	823
841	QY	CACAAGTGAATCTCATACAGAGGCCCTGCGCTTCTTCCACTGTTGCTGACCCCATC	900
824	DB	CACAAGTGAATCTCCATACAGAGGCCCTGCGCTTCTTCCACTGTTGCTGACCCCATC	883
901	QY	CTCTATGCTCTCTCGGGGCCAAGTTCAAAAGCTCTGCCCAGCATGCACTCAACTCCATG	960
884	DB	CTCTAGCCTTCTCTGGGGCCAAATTCAGAAGCTCGCGCAGCATGCACTCAATTCATG	943
961	QY	AGCAGAGGTCGAGCCTCAAGATCTCTTTCCAAAGAAAGCGGGGTGGACACTCTCCGTC	1020
944	DB	AGCAGAGGTCGAGCCTCAAGATCTCTTTCCAAAGGGAAACGGGGTGGACACTCTCCGTC	1003
1021	QY	TCCAGGGAGTCAGATCTCCAGTTTTCATCTCCAGTAACTTATGCAAGACTTATAT	1080
1004	DB	TCCAGAGGTGAGATCTCTCAAGTTTTCATCTCCAGTAACTTATGCAAGACTTATAT	1063
1081	QY	AATATATATATATAT-----GATAAAGAACTTTTTTATGTTACATATTTCCAG	1131
1064	DB	AATATATATATATATATATATGACAAAGAACTTTTTTATGTTACATATTTCCAG	1123
1132	QY	ATATAAGAGCTGACAGCTGTGACAGTTTTTTTTTTTTTAAATGACTGTGGGAGT	1191
1124	DB	ATATAAAGACTGACCAATCTTTGTACAGTTTTTTTTTATGACTGTGGAGTTATGTTTC	1183
1192	QY	TTATGTTCTCT-----AGTTTTGTGAGGTTTGACTTAATTTATATAATAT-----	1239
1184	DB	TCTAGTTTTTGTGTGTGTGTGTGTGTCAGGTTTGACTTAATTTATATAATATTTGGTT	1243
1240	QY	-----TGTTTTTTGTTGTTTCATGTGTAATGAGCGCTTAGCGAGCACTGTGGCCAACT	1293

Db	1244	TTTTCTGTTGTTTGTCTGTTTACGCTGAATGAGTGTCTAGGCAGGACCTGTGTGGCCAAAGT	1303
Qy	1294	TCTTAGTAGCTGTTTATCTCTGTGTAGGACTGTAGAACTGTAGAGGAAGAAACTGAAC	1351
Db	1304	TCCTAGTCGCTGCTGCTGTGTGTAGACTGTAGACTGTAGAGGAAGGACTGAAC	1361
RESULT 15			
AX481810		1582 bp	DNA
LOCUS			linear
DEFINITION		Sequence 11 from Patent WO02057308.	PAT 16-AUG-2002
ACCESSION		AX481810	
VERSION		AX481810.1	GI:22316633
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Danaïson,C., Isalan,M., Moore,M., Choo,Y., Girdlestone,J.R. and Ullman,C.	
TITLE		Nucleic acid binding polypeptides	
JOURNAL		Patent: WO 02057308-A 11 25-JUL-2002;	
FEATURES		SANGAMO BIOSCIENCES INC (US)	
source		Location/Qualifiers	
		1..1582	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		55.1%;	Score 933.4; DB 6; Length 1582;
Best Local Similarity		81.8%;	Pred. No. 2.6e-194;
Matches 1154;		Conservative 0;	Mismatches 216; Indels 41; Gaps 5;
Qy	1	ATATACACTTCTGATAACTACTCTGTGAAGAGTGGGGTCTGGAGACTATGACTCCAAAG	60
Db	7	ATATACACTTCAGATAACTACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG	66
Qy	61	GAACCTGCTTCGGGATGAAAACGTCCTTTCATAGGATCTTCCTGCCACCACCTCTAC	120
Db	67	GAACCTGTTTCGTGAAGAAATGCTAATTTCAATAAATCTCTCTGCCACCACTCTAC	126
Qy	121	TTCAATCATCTCTTGAATGGCATAGTCGGCAATGGATTGGTGATCTCTGGTCAATGGTTAC	180
Db	127	TCCATCATCTCTTAATCGGCATTGTGGGCAATGGATTGGTTCATCTCTGGTCAATGGTTAC	186
Qy	181	CAGAAGAGCTAAGGACATGACGGCAAGTACCGCTGCACCTGTGCAGTGGCTGACCTC	240
Db	187	CAGAAGAACTGAGAAGCATGACGGCAAGTACAGGCTGCACCTGTGCAGTGGCGGACCTC	246
Qy	241	CTCTTTGTCTATCACACTCCCTCTCTGGCCAGTTGATGCCATGGCTGACTGTACTTTGGG	300
Db	247	CTCTTTGTCTATCACGCTTCCTCTCTGGCCAGTTGATGCCGTGGCAACTGGTACTTTGGG	306
Qy	301	AAATTTTGTGTAAAGCTGTCCATATCATCTACACTGTCAACCTCTACAGGAGGTTCTC	360
Db	307	AAC TTCATGCAAGGAGTCCATGTCTATCTACACAGTCAACCTCTACAGAGTGTCTC	366
Qy	361	ATCTGCGCTTCATCAGCCTGGACCGGTGCTCTCGCCATTTGCCACGCCACCAACAGTCAA	420
Db	367	ATCTGCGCTTCATCAGCTCTGGACCGTCTACTTGGCCATCTCTCACGCCACCAACAGTCA	426
Qy	421	AGGCCAAGGAACTGTGGCTGTGAAAAGGAGTCTATGTGGCGTCTGTGATCCAGCCCTC	480
Db	427	AGGCCAAGGAGCTGTGGCTGTGAAAAGGAGTCTATGTGGCGTCTGTGATCCCTCGCCTC	486
Qy	481	CTCTGACTATACCTGACTTTCATCTTTGGCGACCTGCAGCCAGGAGGACATCAGTCAGGG	540
Db	487	CTGCTGACTATTCGCACTTCATCTTTGCCAAGC-----TCAGTAGGCA	531
Qy	541	GATGACAGGTACATCTGTGACCGCTTTATCCCGATAGCTGTGGATGGTGGTCTTTCAA	600

Search completed: May 17, 2004, 12:10:55
Job time : 6860.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 03:46:58 ; Search time 726.44 Seconds
(without alignments)
9906.453 Million cell updates/sec

Title: US-09-367-052-7

Perfect score: 1694

Sequence: 1 atatacactcttgataacta.....taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1694	100.0	1758	2	Aaz27611 Mouse CXC
2	1692.4	99.9	1877	2	Aav46370 Nucleic a
3	933.4	55.1	1582	6	AB564793 CXCR4 rec
4	933.4	55.1	1588	2	Aaz27610 Human CXC
5	933.4	55.1	1664	3	Aaa61656 Human CXC
6	933.4	55.1	1664	3	Aaz40014 CXCR4 cod
7	933.4	55.1	1679	4	Aad12801 Human neu
8	933.4	55.1	1679	4	Aai65467 Nucleotid
9	933.4	55.1	1679	6	AB553392 DNA encod
10	933.4	55.1	1679	6	ABV78039 Hypoxia-r
11	933.4	55.1	1679	7	ABX08779 Angiogene
12	933.4	55.1	1679	7	ABX74454 Human CDN
13	933.4	55.1	1679	7	AB268886 Nucleotid
14	933.4	55.1	1679	7	AB242642 Human CXC
15	933.4	55.1	1679	9	Adc98645 Human CXC
16	933.4	55.1	1711	6	AB235348 Human gen
17	933.4	55.1	1711	9	ADB47320 Human CDN
18	933.4	55.1	1737	2	AAQ80521 Human mon
19	933.4	55.1	1737	2	AAQ99007 Chemokine
20	933.4	55.1	1944	2	AAx15882 cDNA enco
21	933.4	55.1	1944	3	AAx34774 Human ade
22	933.4	55.1	5161	3	Aaf20896 Human CCR
23	933.4	55.1	5161	6	AAH99951 CXCR4 gen

24	933.4	55.1	5161	7	AB296590 Human CCR
25	933.4	55.1	5161	7	ACA64841 Human CXC
26	931.8	55.0	8747	9	ADD67548 Human LY1
27	931.4	55.0	1670	6	ABK83803 Human CDN
28	931.4	55.0	1670	6	ABN95645 Gene #214
29	931.4	55.0	1670	7	ACC46765 Human COP
30	931.4	55.0	1679	6	AB553750 DNA encod
31	930.2	54.9	1737	2	AAQ29506 New plate
32	905.4	53.4	1050	9	AD58421 Toxicity-
33	905.4	53.4	1050	9	AD58421 Primary r
34	855.6	50.5	1225	6	AB235630 Human gen
35	855.6	50.5	1225	7	ACA56637 Human sig
36	843.6	49.8	1224	6	AB235563 Human gen
37	834	49.2	1317	2	AAV18357 Human RM3
38	834	49.2	1317	2	AAV91726 Human 7TM
39	834	49.2	1317	6	ABK54255 Human 7 t
40	832.4	49.1	1317	2	AAQ66179 Seven tra
41	806.2	47.6	1059	3	AAZ38553 Human CXC
42	806.2	47.6	1059	6	ABQ74938 Human CXC
43	806.2	47.6	1059	6	AAH99952 CXCR4 enc
44	804.6	47.5	1059	7	ACC72707 Human can
45	804.6	47.5	1102	7	ACA64750 Chemokine

ALIGNMENTS

RESULT 1
AAZ27611
ID AAZ27611 standard; DNA; 1758 BP.
XX
AC AAZ27611;
XX
DT 16-DEC-1999 (first entry)
DE
DE Mouse CXCR4 coding sequence.
XX CXCR4; mouse; neovascularisation; inhibitor; solid cancer; therapy;
XX tissue repairing agent; vascularisation; ss.
XX Mus sp.
XX WO9948528-A1.
XX
XX 30-SEP-1999.
XX
XX 23-MAR-1999; 99WO-JP001448.
XX
XX 24-MAR-1998; 98JP-00095448.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX (KISH/) KISHIMOTO T.
XX Kishimoto T, Nagasawa T, Tachibana K;
XX WPI; 1999-591042/50.
XX P-PSDB; AAY39994.
XX CXCR4-potentiating agents and methods useful for inhibiting
XX neovascularization, and treating solid cancers.
XX Disclosure; Page 50-51; 63pp; Japanese.
XX This sequence encodes the mouse CXCR4 protein. The invention relates to
XX remedies inhibiting neovascularisation, remedies for solid cancer,
XX remedies for diseases pathologically caused by neovascularisation and
XX tissue repairing agents containing as the active ingredient a substance
XX capable of potentiating CXCR4. Based on a finding that vascularisation is
XX inhibited in a CXCR4 knockout mouse, it becomes possible to prepare
XX remedies inhibiting vascularisation which contain as the active
XX ingredient a substance capable of potentiating CXCR4, remedies for solid
XX cancer, remedies for diseases pathologically caused by neovascularisation
XX and tissue repairing agents containing as the active ingredient a

CC substance capable of potentiating CXCR4. It is also possible to establish
 XX methods for treatment with the use of these remedies

SQ Sequence 1758 BP; 447 A; 414 C; 369 G; 528 T; 0 U; 0 Other;

Query Match 100.0%; Score 1694; DB 2; Length 1758;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1694; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAAAG 60
 Db 22 ATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCAAAG 81
 Qy 61 GAACCCCTCTCCGGGATGAAGCTCCATTTCAATAGGATCTTCCGCCACCACTAC 120
 Db 82 GAACCCCTCTCCGGGATGAAGCTCCATTTCAATAGGATCTTCCGCCACCACTAC 141
 Qy 121 TTATCATCTCTTCTGACTGGCATAGTGGCAATGGATTTGGTGTGATCTCTGGTTCAC 180
 Db 142 TTATCATCTCTTCTGACTGGCATAGTGGCAATGGATTTGGTGTGATCTCTGGTTCAC 201
 Qy 181 CAGAAGAGCTAAGAGGATGACGGACAAAGTACCGGTGACCTGTGAGTGGGTGACCTC 240
 Db 202 CAGAAGAGCTAAGAGGATGACGGACAAAGTACCGGTGACCTGTGAGTGGGTGACCTC 261
 Qy 241 CTCTTTGTGTCACACTCCCTCTCTGGCAGTGTGATGCCATGGCTGACTGTACTTTGGG 300
 Db 262 CTCTTTGTGTCACACTCCCTCTCTGGCAGTGTGATGCCATGGCTGACTGTACTTTGGG 321
 Qy 301 AAATTTTGTGTAGGCTGTCCATATCATCTACACTGTCTCAACCTCTACAGAGGTTCTC 360
 Db 322 AAATTTTGTGTAGGCTGTCCATATCATCTACACTGTCTCAACCTCTACAGAGGTTCTC 381
 Qy 361 ATCTGGCTTTCATCAGCTCGGACGGTGTCTCCGCTATGTCACGCGCCACCAAGTCAA 420
 Db 382 ATCTGGCTTTCATCAGCTCGGACGGTGTCTCCGCTATGTCACGCGCCACCAAGTCAA 441
 Qy 421 AGGCCAAGAAACTGCTGGCTGAAAGGCACTGTATGTGGCGCTCTGGATGCCAGCCCTC 480
 Db 442 AGGCCAAGAAACTGCTGGCTGAAAGGCACTGTATGTGGCGCTCTGGATGCCAGCCCTC 501
 Qy 481 CTCTGACTATACCTGACTTCTATCTTTGCGACGCTCAGCCAGGGGGACATCAGTCAGGG 540
 Db 502 CTCTGACTATACCTGACTTCTATCTTTGCGACGCTCAGCCAGGGGGACATCAGTCAGGG 561
 Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA 600
 Db 562 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCCTGTGGATGGTGGTGTTCAA 621
 Qy 601 TTCAGCATATATGGTGGGTCTCATCTGCGCGGATCTGTCATCTCTCTGTACTGTC 660
 Db 622 TTCAGCATATATGGTGGGTCTCATCTGCGCGGATCTGTCATCTCTCTGTACTGTC 681
 Qy 661 ATCATCATCTCTAAGCTGTCACTTCCAGGGGCCACAGAGCGCAAGGCCCTCAAGACG 720
 Db 682 ATCATCATCTCTAAGCTGTCACTTCCAGGGGCCACAGAGCGCAAGGCCCTCAAGACG 741
 Qy 721 ACAGTCATCTCATCTAGCTTTCTTTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780
 Db 742 ACAGTCATCTCATCTAGCTTTCTTTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 801
 Qy 781 ATCGACTCTTCATCTTTTGGAGTGTATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG 840
 Db 802 ATCGACTCTTCATCTTTTGGAGTGTATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG 861
 Qy 841 CACAAGTGGATCTCCATCACAAGAGCCCTCGCCCTCTTCCACTGTTCGCTGAACCCCATC 900
 Db 862 CACAAGTGGATCTCCATCACAAGAGCCCTCGCCCTCTTCCACTGTTCGCTGAACCCCATC 921
 Qy 901 CTCTATGCTCTCTCGGGGCCAAGTTCAAGAGCTCTGCCAGCATGCACTCACTCCCATG 960
 Db 922 CTCTATGCTCTCTCGGGGCCAAGTTCAAGAGCTCTGCCAGCATGCACTCACTCCCATG 981

Qy 961 AGCAGAGCTCCAGCTCAAGATCTTTTCCAAAGAAAGCGGGTGGACACTCTTCGCTC 1020
 Db 982 AGCAGAGCTCCAGCTCAAGATCTTTTCCAAAGAAAGCGGGTGGACACTCTTCGCTC 1041
 Qy 1021 TCCACGGAGTCAAGATCTCTCAGTCTTTCACCTCCAGCTAAACCTTATGCAAGACTTAT 1080
 Db 1042 TCCACGGAGTCAAGATCTCTCAGTCTTTCACCTCCAGCTAAACCTTATGCAAGACTTAT 1101
 Qy 1081 AAT 1140
 Db 1102 AAT 1161
 Qy 1141 ACTGACCACTCTGTGACAGTCTTTTCTTTTAAATTAATTAATTAATTAATTAATTA 1200
 Db 1162 ACTGACCACTCTGTGACAGTCTTTTCTTTTAAATTAATTAATTAATTAATTAATTA 1221
 Qy 1201 TCTAGTCTTTGTAGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
 Db 1222 TCTAGTCTTTGTAGGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1281
 Qy 1261 GAATGAGCGTCTAGGACGACCTGTGGCCAAAGTCTTGTAGTGTCTTATCTGTGTAG 1320
 Db 1282 GAATGAGCGTCTAGGACGACCTGTGGCCAAAGTCTTGTAGTGTCTTATCTGTGTAG 1341
 Qy 1321 GACTGTAGAACTGTAGAGGAAAGAACTGAACATTCAGAAATGTGTGTAATTTGAATAA 1380
 Db 1342 GACTGTAGAACTGTAGAGGAAAGAACTGAACATTCAGAAATGTGTGTAATTTGAATAA 1401
 Qy 1381 GCTAGCGGTGATCTCAGCTGTGTGTCATATCTCTTCATCCGAGGACACCCACCC 1440
 Db 1402 GCTAGCGGTGATCTCAGCTGTGTGTCATATCTCTTCATCCGAGGACACCCACCC 1461
 Qy 1441 CCACCCCAACCCCAACCCCACTTCTTAAATTTGTTGTTATGCTGTGTGTGTTGTTG 1500
 Db 1462 CCACCCCAACCCCAACCCCACTTCTTAAATTTGTTGTTATGCTGTGTGTGTTGTTG 1521
 Qy 1501 GTTTTGT 1560
 Db 1522 GTTTTGT 1581
 Qy 1561 GCTGAAATGT 1620
 Db 1582 GCTGAAATGT 1641
 Qy 1621 AGAGTAGATGTAGTCTCAGTCCCTACAAATGTACAGTCTTGTATTAACATGTTAAAG 1680
 Db 1642 AGAGTAGATGTAGTCTCAGTCCCTACAAATGTACAGTCTTGTATTAACATGTTAAAG 1701
 Qy 1681 TCAATGATTAACCTT 1694
 Db 1702 TCAATGATTAACCTT 1715

RESULT 2

AAV46370

ID AAV46370 standard; cDNA to mRNA; 1877 BP.

XX AAV46370;

XX 20-NOV-1998 (first entry)

XX Nucleic acid encoding a murine CXC chemokine receptor.

XX Mouse; CXC chemokine receptor; pre-B cell line DW34;

XX CXC chemokine pre-B cell stimulatory factor PSF/SDF-1; HIV infection;
 screening; inhibitor; AIDS; ds.

XX Mus sp.

XX Key Location/Qualifiers

XX CDS 120..1199

XX FT /*tag= a

XX XX

PN WO9835035-A1.
 XX 13-AUG-1998.
 XX 07-FEB-1997; 97WO-JP000299.
 XX 07-FEB-1997; 97WO-JP000299.
 XX (SHIO) SHIONOGI & CO LTD.
 XX Kishimoto T, Nagasawa T, Tachibana K, Iizasa H, Yoshida N;
 PI Nakajima T, Yoshie O;
 XX P-PSDB; AA64778.
 DR WPI: 1998-447232/38.
 XX P-PSDB; AA64778.
 XX Mouse CXK chemokine receptor binding to PBSF/SDF-1 pre-B cell stimulatory
 PT factor - is useful for screening of potential HIV infection and AIDS
 PT inhibitors.
 XX Claim 3; Page 39-42; 76pp; Japanese.
 XX The present sequence encodes a murine CXK chemokine receptor which binds
 CC to the mouse CXK chemokine pre-B cell stimulatory factor PBSF/SDF-1. The
 CC nucleic acid is isolated from mouse pre-B cell line DW34. The receptor
 CC and cells expressing it can be used in the study and mapping of the
 CC mechanism of HIV infection and in screening of potential inhibitors of
 CC HIV infection and the development of AIDS
 XX Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T; 0 U; 0 Other;
 SQ Query Match 99.9%; Score 1692.4; DB 2; Length 1877;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATATACACTTCTGATACTCTGAGAGAGTGGGCTGGAGACTATGACTCCACAG 60
 DB 141 ATATACACTTCTGATACTCTGAGAGAGTGGGCTGGAGACTATGACTCCACAG 200
 QY 61 GAACCTGCTTCGGGATGAAACGTCATTTCAATAGATCTTCCGCCACCATCTAC 120
 DB 201 GAACCTGCTTCGGGATGAAACGTCATTTCAATAGATCTTCCGCCACCATCTAC 260
 QY 121 TTCAATCTTCTGACTGGCATAGTGGCAATGGAATGGTATCCTGGTCATGGGTAC 180
 DB 261 TTCAATCTTCTGACTGGCATAGTGGCAATGGAATGGTATCCTGGTCATGGGTAC 320
 QY 181 CAGAGAGCTAAGGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 240
 DB 321 CAGAGAGCTAAGGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGCTGACCTC 380
 QY 241 CTCCTTGTCTACACTCCCTTCTGGGAGTGTGATGCCATGGCTGACTGGTACTTTGGG 300
 DB 381 CTCCTTGTCTACACTCCCTTCTGGGAGTGTGATGCCATGGCTGACTGGTACTTTGGG 440
 QY 301 AAATTTTGTAGGCTGTCCATATCATCTACATGTCMACTCTACAGCAGCGTCTC 360
 DB 441 AAATTTTGTAGGCTGTCCATATCATCTACATGTCMACTCTACAGCAGCGTCTC 500
 QY 361 ATCCTGGCTTCATAGCCTGACCGGTACCTCGGCATTTGCCAGCCACCAACAGTCAA 420
 DB 501 ATCCTGGCTTCATAGCCTGACCGGTACCTCGGCATTTGCCAGCCACCAACAGTCAA 560
 QY 421 AGCCCAAGAAACTGTGGCTGAAAGGAGCTATGTGGGGCTGTGATCCAGCCCTC 480
 DB 561 AGCCCAAGAAACTGTGGCTGAAAGGAGCTATGTGGGGCTGTGATCCAGCCCTC 620
 QY 481 CTCCTGACTATACCTGACTTCTTTTGGCGAGTCAGCCAGGGGACATCAGTCAGGGG 540
 DB 621 CTCCTGACTATACCTGACTTCTTTTGGCGAGTCAGCCAGGGGACATCAGTCAGGGG 680
 QY 541 GATGACAGGTACATCTGTGACCGCTTTTACCCCGATAGCCTGTGAGATGGTGTGTTCAA 600

DB 581 GATGACAGGTACATCTGTGACCGCTTTTACCCTGATAGCCTGTGGATGGTGGTTTCAA 740
 QY 601 TTCCAGCATATATATGTGGGTCTCATCTGCGCGGCAATGTCATCTCTCTGTTTACTGC 860
 DB 741 TTCCAGCATATATATGTGGGTCTCATCTGCGCGGCAATGTCATCTCTCTGTTTACTGC 800
 QY 661 ATCATCATCTTAAGCTGTACACTCCAAAGGGCCACCAAGCGCAAGCCCTCAAGACG 720
 DB 801 ATCATCATCTTAAGCTGTACACTCCAAAGGGCCACCAAGCGCAAGCCCTCAAGACG 860
 QY 721 ACAGTTCATCT 780
 DB 861 ACAGTTCATCT 920
 QY 781 ATCGACT 840
 DB 921 ATCGACT 980
 QY 841 CACAAGTGGATCTCCATCATCAGAGGCGCTCGCCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 DB 981 CACAAGTGGATCTCCATCATCAGAGGCGCTCGCCTCTCTCTCTCTCTCTCTCTCTCTCT 1040
 QY 901 CTCTATGCT 960
 DB 1041 CTCTATGCT 1100
 QY 961 AGCAGAGGCTCCAGCTCTCAAGATCTCTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTC 1020
 DB 1101 AGCAGAGGCTCCAGCTCTCAAGATCTCTTCCAAAGGAAAGCGGGTGGACACTCTTCCGTC 1160
 QY 1021 TCCAGGAGTCCAGATCT 1080
 DB 1161 TCCAGGAGTCCAGATCT 1220
 QY 1081 AAT 1140
 DB 1221 AAT 1280
 QY 1141 ACTGACAGCTTTGTACAGTT 1200
 DB 1281 ACTGACAGCTTTGTACAGTT 1340
 QY 1201 TCTAGTTTTTGTAGGTTTGTACCTTAATATATATATATATATATATATATATATATAT 1260
 DB 1341 TCTAGTTTTTGTAGGTTTGTACCTTAATATATATATATATATATATATATATATATAT 1400
 QY 1261 GAATGAGCGTCTAGGAGGACCTGTGGCCAAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
 DB 1401 GAATGAGCGTCTAGGAGGACCTGTGGCCAAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1460
 QY 1321 GACTGTAGAACTGTAGAGGAGAAACTGAACATTCAGATGTGTGTAAATGAATATAA 1380
 DB 1461 GACTGTAGAACTGTAGAGGAGAAACTGAACATTCAGATGTGTGTAAATGAATATAA 1520
 QY 1381 GCTAGCGCTGATCTCAGCTGTGTGCTGCAATCTCTTCAATCCGAGGAGCACCCACCC 1440
 DB 1521 GCTAGCGCTGATCTCAGCTGTGTGCTGCAATCTCTTCAATCCGAGGAGCACCCACCC 1580
 QY 1441 CCACCCCCACCCCAACCCCAATCTTAAATGTTTTGTTATGCTGTGTGTGTGTGTGTGT 1500
 DB 1581 CCACCCCCACCCCAACCCCAATCTTAAATGTTTTGTTATGCTGTGTGTGTGTGTGTGT 1640
 QY 1501 GTTTTTTTTTTGT 1560
 DB 1641 GTTTTTTTTTTGT 1700
 QY 1561 GCCTGAAATGGTGGTGAATAAGTCTGGGGTTTTTTTTTTTTTTTTTTTTTTTTTCAAGTTTCA 1620
 DB 1701 GCCTGAAATGGTGGTGAATAAGTCTGGGGTTTTTTTTTTTTTTTTTTTTTTTTTCAAGTTTCA 1760
 QY 1621 AGAGTAGATTGACTTCAAGTCCCTACAAATGTACAGCTTGTATATATATATATATATAT 1680
 DB 1761 AGAGTAGATTGACTTCAAGTCCCTACAAATGTACAGCTTGTATATATATATATATATAT 1820

Db 1064 GACTTTTATACGATAAATAAATACTTTTTPAAGTTACACATTTTTCAGATATAAAAG 1123
Qy 1141 ACTGACAGCTCTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGTTGGGAGTTTATGTTCC 1200
Db 1124 ACTGACCAATATGTACAGTTTTTATTTGCTGTGGATTTTGT-----CTTGTTTC 1177
Qy 1201 TCTAGTTTTGTGAGGTTTGTGACTTAATTTATATAAATATGTTTTTGTGTTTCATGT 1260
Db 1178 TTTAGTTTTTGTGAAGTTTAAATGACTTAATTTATA-----TAAATTTTTTTTGTTCATAT 1233
Qy 1261 GAATGAGCGTCTAGGAGGACCTGTGGCAGATTTCTTAGTAGTGTGTTTATCTGTGTGTAG 1320
Db 1234 TGATGTGTCTAGGAGGACCTGTGGCAGATTTCTTAGTAGTGTGTTTATCTGTGTGTAG 1293
Qy 1321 GACTGTAGAACTGTAGAGGAGAAATGAACTCCAGAAATGTTGGTAAATTTGAATAAA 1380
Db 1294 GACTGTAGAA-----AAGGGAATGAACTCCAGAGCGGTAGTGAATCACTCACTAAA 1345
Qy 1381 GCTAGCGCTGATCCTCAGCTGTTGCTGCATA 1411
Db 1346 GCTAGAAATGATCCAGCTGTTTATGCTATA 1376

RESULT 4

AAZ27610

ID AAZ27610 standard; DNA; 1588 BP.

XX AC AAZ27610;

XX DT 16-DEC-1999 (first entry)

XX DE Human CXCR4 coding sequence.

XX KW CXCR4; human; neovascularisation; inhibitor; solid cancer; therapy;

XX KW tissue repairing agent; vascularisation; ss.

XX OS Homo sapiens.

XX PN WO948528-A1.

XX PD 30-SEP-1999.

XX PF 23-MAR-1999; 99WO-JP001448.

XX PR 24-MAR-1998; 98JP-00095448.

XX PA (CHUS) CHUGAI SEIYAKU KK.

XX PA (KISH) KISHIMOTO T.

XX PI Kishimoto T, Nagasawa T, Tachibana K;

XX DR WPI; 1999-591042/50.

XX DR P-PSDB; AAY39993.

XX CXCR4-potentiating agents and methods useful for inhibiting

PT neovascularization, and treating solid cancers.

XX Disclosure; Page 48-49; 63pp; Japanese.

XX This sequence encodes the human CXCR4 protein. The invention relates to

CC remedies inhibiting neovascularisation, remedies for solid cancer,

CC remedies for diseases pathologically caused by neovascularisation and

CC tissue repairing agents containing as the active ingredient a substance

CC capable of potentiating CXCR4. Based on a finding that vascularisation is

CC inhibited in a CXCR4 knockout mouse, it becomes possible to prepare

CC remedies inhibiting vascularisation which contain as the active

CC ingredient a substance capable of potentiating CXCR4, remedies for solid

CC cancer, remedies for diseases pathologically caused by neovascularisation

CC and tissue repairing agents containing as the active ingredient a

CC substance capable of potentiating CXCR4. It is also possible to establish

XX methods for treatment with the use of these remedies

XX Sequence 1588 BP; 396 A; 375 C; 334 G; 483 T; 0 U; 0 Other;

Query Match 55.1%; Score 933.4; DB 2; Length 1588;
Best Local Similarity 81.8%; Pred. No. 4.2e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGTTCTGGAGACTATGATCTCAACAAG 60
Db 16 ATATACACTTCTGATACTACTCTGAAGAAGTGGGTTCTGGAGACTATGATCTCAACAAG 75
Qy 61 GAAACCTCTCTCCGGGATGAAAACGTCATTTCAATAGGATCTTCTCTCCGCCACCATCTAC 120
Db 76 GAAACCTCTCTCCGGGATGAAAACGTCATTTCAATAGGATCTTCTCTCCGCCACCATCTAC 135
Qy 121 TCTATCATCTTCTTGTGACTGCGCATAGTCGGCAATGGATTTGGTATCTCTGGTTCATGGTTAC 180
Db 136 TCCATCATCTTCTTAACTGGCAATTTGGGCAATGGATTTGGTATCTCTGGTTCATGGTTAC 195
Qy 181 CAGAAGAAGCTTAAGGAGCATGACGACCAAGTACCGGCTGCACCTGTGTCAGTGGCTGACCTC 240
Db 196 CAGAAGAAGCTTAAGGAGCATGACGACCAAGTACCGGCTGCACCTGTGTCAGTGGCTGACCTC 255
Qy 241 CTCCTTGTATCATCACATCCCTTCTGGGCAAGTTGATGCCATGGCTGATGCTATCTTTGGG 300
Db 256 CTCCTTGTATCATCACATCCCTTCTGGGCAAGTTGATGCCATGGCTGATGCTATCTTTGGG 315
Qy 301 AAATTTTGTGTAAGGCTGCCATATCATCTACACTGTCAACCTCTACAGCAGGTTCTC 360
Db 316 AAATTTTGTGTAAGGCTGCCATATCATCTACACTGTCAACCTCTACAGCAGGTTCTC 375
Qy 361 ATCTGGCCTTTCATCAGCTGGACCGGTACCTGCCATTTGTCCACGCCACCAACAGTCAA 420
Db 376 ATCTGGCCTTTCATCAGCTGGACCGGTACCTGCCATTTGTCCACGCCACCAACAGTCAA 435
Qy 421 AGGCCAAGGAAGCTGTGGCTGAAAAGGCACTGATGTGGGCTCTGGATCCCAAGCCCTC 480
Db 436 AGGCCAAGGAAGCTGTGGCTGAAAAGGCACTGATGTGGGCTCTGGATCCCAAGCCCTC 495
Qy 481 CTCCTGACTATACCTGACTTTCATCTTTGCCGACCTCAGCCAGGGGGACATCAGTCAGGGG 540
Db 496 CTCCTGACTATACCTGACTTTCATCTTTGCCGACCTCAGCCAGGGGGACATCAGTCAGGGG 540
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTCAC 600
Db 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTCAC 600
Qy 601 TTCAGCATATAATGGTGGGTCATCTCTGCCGCGCATCGTCATCTCTCTCTCTCTCTCTG 660
Db 601 TTCAGCATATAATGGTGGGTCATCTCTGCCGCGCATCGTCATCTCTCTCTCTCTCTCTG 660
Qy 661 ATCATCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAAGCGCAAGGGCCCTCAAGACG 720
Db 661 ATCATCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAAGCGCAAGGGCCCTCAAGACG 720
Qy 721 ACAGTCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAAGCGCAAGGGCCCTCAAGACG 780
Db 721 ACAGTCATCTCTAAGCTGTCACTTCCAAAGGGCCACCAAGCGCAAGGGCCCTCAAGACG 780
Qy 781 ATCGACTCTTTCATCTCTTGGGAGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG 840
Db 781 ATCGACTCTTTCATCTCTTGGGAGTTCATCAAGCAAGGATGTGACTTCGAGAGCATTTGTG 840
Qy 841 CACAAGTGGATCTCCATCAGAGGCGCTGCTTCTTCCATCTGTGCTGCTGAAACCCATC 900
Db 841 CACAAGTGGATCTCCATCAGAGGCGCTGCTTCTTCCATCTGTGCTGCTGAAACCCATC 900
Qy 901 CTCCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTCTGCCAGCATGCACTCAACTCCATG 960
Db 901 CTCCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTCTGCCAGCATGCACTCAACTCCATG 960
Qy 961 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAGCGGGGTGAGCACTCTTCCGTC 1020
Db 961 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAGCGGGGTGAGCACTCTTCCGTC 1020

QY 1021 TCCAGGAGTCAGATCCTCCAGTTTTCACCTCCAGCTAACCCCTATGCAAGACTTATAT 1080
Db 1021 TCCAGTGAAGTCTGAGTCTTCAAGTTTTCACCTCCAGCTAA-----CACAGATGTAAA 1072
QY 1081 AATATATATATATATATATGATAAAGAACTTTTATGTTACACATTTTCCAGATATAAG 1140
Db 1073 GACTTTTTTATAGATAAATAACTTTTTTAAAGTTACACATTTTTCAGATATAAG 1132
QY 1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCC 1200
Db 1133 ACTGACCAATATGTACAGTTTTTATGCTGTGTTGGATTTTGT-----CTTGTGTTTC 1186
QY 1201 TCTAGTTTTTGTGAGTTTGTACTTAATTTATATAAATATTGTTTTTGTGTTTTCATGT 1260
Db 1187 TTTAGTTTTTGTGAGTTTAAATGACTATTATATA-----TAAATTTTTTGTGTTTCATAT 1242
QY 1261 GAATGAGCGTCTAGCAGAGACCTGTGGCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
Db 1243 TGATGTGTGTCTAGCAGAGACCTGTGGCAAGTTCTTAGTTGCTGTATGCTGTGTGTAG 1302
QY 1321 GACTGTAGAACTGTAGAGGAAGAACTGACATCCAGATGTGTGTTAAATGATATAA 1380
Db 1303 GACTGTAGAA-----AAGGGAAGTGAACATTTCCAGAGCGTGTAGTTAATCACGTAA 1354
QY 1381 GCTAGCGGTGATCTCAGCTTTCTGTCATA 1411
Db 1355 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1385

RESULT 5

AAAG61656
ID AAAG61656 standard; DNA; 1664 BP.
AC AAAG61656;
XX
DT 23-OCT-2000 (first entry)
XX Human CXCR4 chemokine receptor (CXCR4) cDNA.
DE CXCR4; human CXCR4 chemokine receptor 4; HIV infection inhibition;
XX expression inhibition; antisense therapy; ss.
KW Homo sapiens.
XX
FH Key Location/Qualifiers
FT primer_bind 61..83
FT /*tag= b
FT /bound_moiety= "Antisense oligonucleotide SEQ ID NO:3
FT (AAAG61654)"
FT primer_bind 67..90
FT /*tag= a
FT /bound_moiety= "Antisense oligonucleotide SEQ ID NO:1
FT (AAAG61652)"
FT primer_bind 73..96
FT /*tag= b
FT /bound_moiety= "Antisense oligonucleotide SEQ ID NO:2
FT (AAAG61653)"
XX
FN WO2000031271-A1.
XX
PD 02-JUN-2000.
XX
PF 24-NOV-1999; 99WO-JP006534.
XX
PR 24-NOV-1998; 98JP-00332760.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
XX Iijima O, Goto T, Shimada T;
XX WPI; 2000-400081/34.
XX Antisense oligonucleotide, useful as inhibitor in preventive or remedy

PT for HIV infection with high therapeutic efficacy.
XX Example 1; Page 17-18; 21pp; Japanese.
XX
CC The invention relates to human CXCR4 chemokine receptor 4 (CXCR4)-specific
CC antisense oligonucleotides (AAAG61652-AA61654). The oligonucleotides
CC hybridize with the CXCR4 gene and/or CXCR4 RNA, thereby preventing its
CC expression. As CXCR4 plays a role in HIV infection, the antisense
CC oligonucleotides act as inhibitors of this process. Compositions
CC containing the antisense oligonucleotides are useful as prophylactic or
CC therapeutic agents for HIV infection. Such compositions are highly
CC efficacious for inhibiting infection with HIV. The present sequence
CC represents a human CXCR4 cDNA
XX
SQ Sequence 1664 BP; 414 A; 398 C; 359 G; 493 T; 0 U; 0 Other;
Query Match 55.1%; Score 933.4; DB 3; Length 1664;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACATCTTCGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGATCCCAACAG 60
Db 92 ATATACATCTTCGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGATCCCAACAG 151
QY 61 GAACCTCTGCTTCCGGGATGAAAACGTCATTTTCAATAGGATCTTCTGCCACCATCTTAC 120
Db 152 GAACCTCTGCTTCCGGGATGAAAACGTCATTTTCAATAGGATCTTCTGCCACCATCTTAC 211
QY 121 TTCATCATCTTCTTACCTGGCATAGTCGGCAATGGAATGGATGGTATGATGGGTATAC 180
Db 212 TCCATCATCTTCTTAACTGGCATATGTGGCAATGGAATGGTATGATGGGTATAC 271
QY 181 CAGAAGAAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTGTAGTGGCTGACCTC 240
Db 272 CAGAAGAAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTGTAGTGGCTGACCTC 331
QY 241 CTCTTTGTATGACACTCCCTTCTGGGAGTGTGATGCCATGGCTGACTGTACTTTGGG 300
Db 332 CTCTTTGTATGACACTCCCTTCTGGGAGTGTGATGCCATGGCTGACTGTACTTTGGG 391
QY 301 AAATTTTGTGTAGGCTGTCCATATCATCTACATGTCAACCTCTACAGCAGCGTTCTC 360
Db 392 AACTTCTATGCAAGGAGTGTCAATGTATCTACAGTCAACCTCTACAGCAGTGTCTC 451
QY 361 ATCTGGCTTTCATCAGCTGGACCGGTACCTGCGCATTTGCCAGCCACCAAGTCAAA 420
Db 452 ATCTGGCTTTCATCAGCTGGACCGGTACCTGCGCATTTGCCAGCCACCAAGTCAAA 511
QY 421 AGGCCAAGGAAGTCTGGCTGAAAGGAGTCTATGTGGGCTGTGATCCGAGCTC 480
Db 512 AGGCCAAGGAAGTCTGGCTGAAAGGAGTCTATGTGGGCTGTGATCCGAGCTC 571
QY 481 CTCTGTACTATACCTGACTTTCATCTTCCGACCTGACCCAGGGGACATCAGTCAGGG 540
Db 572 CTGTGTACTATTCGCGACTTTCATCTTTGCCAAG-----TCAGTGAGGCA 616
QY 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGATGGTGGTGTTCAA 600
Db 617 GATGACAGGTATATCTGTGACCGCTTTACCCGATAGCTGTGATGGTGGTGTTCAG 676
QY 601 TTCCAGCATATAATGGTGGGTCTCATCTGCGCGCATCTGTCATCTCTCTGTACTGCG 660
Db 677 TTTGACGACATCATGTGTTGGCTTATCTGCTGTGTTGTTGTCATCTCTCTGTACTGCG 736
QY 661 ATCATCATCTTATAGCTGTCTACCTCCAGGGCCACCGAAGCGCAAGGCCCTCAAGACG 720
Db 737 ATTATCATCTCCAAAGCTGTCTACCTCCAAAGGGCCACCGAAGCGCAAGGCCCTCAAGAC 796
QY 721 ACAGTCATCTCATCTAGCTTTCTTGTGCTGTGGTGTGCTATTTATGTGGGATCAGC 780
Db 797 ACAGTCATCTCATCTAGCTTTCTTGTGCTGTGGTGTGCTATTTATGTGGGATCAGC 856
QY 781 ATCGACTCTTTCATCTCTTTTGGGAGTCAATCAAGCAAGATGTGACTTTCGAGAGCATGTG 840

QY 841 CACAAGTGGATCTCCATCAGAGAGCCCTCGCTTCTTCCACTGTTGCTGACCCCATC 900
Db 917 CACAAGTGGATCTCCATCAGAGAGCCCTAGCTTCTTCCACTGTTGCTGACCCCATC 976
QY 901 CTCATGCTCTCTCGGGGCAAGTTCAAAAGCTGTCGCCAGCATCACTCACTCCATG 960
Db 977 CTCATGCTCTCTCGGGGCAAGTTCAAAAGCTGTCGCCAGCATCACTCACTCTG 1036
QY 961 AGCAGAGGCTCCAGCTCAGATCTCTTCCAAAGGAAAGCGGGTGGACACTCTCCGTC 1020
Db 1037 AGCAGAGGCTCCAGCTCAGATCTCTTCCAAAGGAAAGCGGGTGGACACTCTCTGTT 1096
QY 1021 TCCAGGAGTCCAGATCTCTCCAGTCTTCACTCCAGCTTAACCTTATGCCAAGACTTATAT 1080
Db 1097 TCCAGTGGTCTGAGTCTTCAAGTCTTCACTCCAGCTAA-----CACAGATGTAATA 1148
QY 1081 AATATATATATATATGTAAGAACTTTTATGTTACACATTTTCAGATATAGAG 1140
Db 1149 GACTTTTTTTTATACGATTAATACTTTTTTTTAAAGTTACACATTTTCAGATATAGAG 1208
QY 1141 ACTGACGAGTCTGTACAGTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCC 1200
Db 1209 ACTGACCAATATGACAGTTTTTATGTTGTTGGATTTTGT-----CTGTGTTTC 1262
QY 1201 TCTAGTTTTTGTGAGGTTGACTTAATTTATATAAATATGTTTTTTTGTGTTGTTCACT 1260
Db 1263 TTTAGTTTTTGTGAGGTTTAAATGACTTAATTTATA-----TAAATTTTTTTTGTGTTCACTAT 1318
QY 1261 GAATGAGCGTCTAGGAGGAGCTGTGCCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
Db 1319 TGATGTGTCTAGGAGGAGCTGTGCCAAGTTCTTAGTTGCTGTATGTTCTGTGTAG 1378
QY 1321 GACTGTAGAACTGTAGAGGAGAACTGAACATTCAGAGATGTGTGTAATTTGAATATA 1380
Db 1379 GACTGTAGAA-----AAGGGAAGTGAACATTCAGAGCGTGTAGTTAATCACGTAAA 1430
QY 1381 GCTAGCGTGTATCTCAGCTGTTGCTGCATA 1411
Db 1431 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1461

RESULT 7
AAD12801
ID AAD12801 standard; DNA; 1679 BP.
XX
AC AAD12801;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human neurotrophin Y (NPY) Y3 receptor DNA.
XX
KW Human; neurotrophin Y; NPY; bone disease; bone mass; gene therapy;
KW cerebrospinal fluid; CSF; inositol phosphate; IP; Paget's disease;
KW fracture; extracellular signal-regulated kinase; ERK; osteoporosis;
KW osteopenia; bone metastasis; neurotransmitter; osteogenic;
KW NPY Y3 receptor; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 89..1147
FT /*tag= a
FT /product= "Human neurotrophin Y (NPY) Y3 receptor"
XX
PN W0200153477-A1.
XX
PD 26-JUL-2001.
XX
PF 22-JAN-2001; 2001WO-US002040.
XX
PR 20-JAN-2000; 2000US-00489872.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.

(AMLI/) AMLING M.
PI Amling M, Karsenty G, Ducey P;
XX
WPI; 2001-488709/53.
DR P-PSDB; AAE06690.
XX
PT Treating or preventing reduced bone mass, e.g. osteoporosis, by reducing
the level of neurotrophin Y activity in blood or cerebrospinal fluid.
XX
PS Example 7; Page 84-85; 102pp; English.
XX
CC The present invention relates to a method for treatment or prevention of
bone diseases characterized by loss of bone mass, comprises administering
to a mammal a compound that lowers the level of neurotrophin Y (NPY) in
the serum or cerebrospinal fluid (CSF) or a compound that lowers the
level of inositol phosphate (IP) or extracellular signal-regulated kinase
(ERK). The method is specifically used to treat (including by gene
therapy) or prevent osteoporosis, osteopenia or Paget's disease. But may
also be used e.g., in cases of fractures or bone metastases. These
CC diseases may also be diagnosed by detecting elevated NPY levels.
CC including monitoring of treatment, assessing efficacy of compounds in
clinical trials and for identifying subjects at risk. The present
sequence is a human NPY Y3 receptor DNA
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 55.1%; Score 933.4; DB 4; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACATCTCTGATACTACTCTGAAGAGTGGGGTCTGGAGACTATGATCTCAACAG 60
Db 104 ATATACATCTCTGATACTACTACCGGAGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163
QY 61 GAACCTCTCTTCCGGGATGAAACGTCATTTCAATAGGATCTTCTGCCCCACCATCTAC 120
Db 164 GAACCTCTCTTCCGGGAGAAATGCTAATTTCAATAAATCTTCTGCCCCACCATCTAC 223
QY 121 TCTATCATCTTCTTGTGACTGTGGCATAGTCGGCAATGGATTTGTGATCTCGGTTCATGGGTAC 180
Db 224 TCCATCATCTTCTTAACTGGCATTTGGGCAATGGATTTGTGATCTCGGTTCATGGGTAC 283
QY 181 CAGAAGAGCTTAAGGAGCATGACGAGCAAGTACCGGCTGCACCTGTGAGTGGGTGACCTC 240
Db 284 CAGAAGAACTGAGAAGCATGACGAGCAAGTACAGGCTGCACCTGTGAGTGGGTGACCTC 343
QY 241 CTCTTTGTGATCACAACCTCCCTTCTGGGCGAGTTGATGCCATGGCTGATCTGTTGTTGGG 300
Db 344 CTCTTTGTGATCACAACCTCCCTTCTGGGCGAGTTGATGCCGAGCAAACTGTTGTTGGG 403
QY 301 AAATTTTGTGTAAGGCTGTCCATATCATCTACTGTCAACCTCTACAGCAGGTTCTC 360
Db 404 AACTTCTTATGCAAGGCGAGTCCATGTCTATCTACACAGTCAACCTCTACAGCAGGTTCTC 463
QY 361 ATCTGGGCTTTCATCAGCTGGAGCGGTACCTGCCATTTGCCACGCCCAACAGTCAA 420
Db 464 ATCTGGGCTTTCATCAGCTGGAGCGGTACCTGCCATTTGCCACGCCCAACAGTCAA 523
QY 421 AGGCCAAGAACTGTGCTGTTGTAAGGAGGAGTCTATGTGGCGTCTGATCTCCAGCCCTC 480
Db 524 AGGCCAAGAACTGTGCTGTTGTAAGGAGGAGTCTATGTGTGGCGTCTGATCTCCCTGCCCTC 583
QY 481 CTCTGTGATACCTGACTGATCTTCTTGGCGAGCTCAGCCAGCGGAGCATCAGTCAAGGG 540
Db 584 CTGTGTGATATCTCCGACTTCACTTTTGGCAGG-----TCAGTGAAGCA 628
QY 541 GATGACAGGTATCTGTGACCGCTTTTACCCGATAGCTGTGGATGGTGTGTTCAA 600
Db 629 GATGACAGATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGTGTTGTTCCAG 688
QY 601 TTCAGCATATATGTTGGGTCTCATCTGCGCGGATCTCATCTCTCTCTCTCTCTCTCTCT 660

Db 689 TTTGAGCAGCATCATGTTGGGCTTATCTCTGCTGGTATTGTTCATCTGCTCTATTGC 748
QY 661 ATCATCATCTCTAAGCTGTACACTCCAGGGCCACAGAGCGGAGCCCTCAAGAGG 720
Db 749 ATTATCATCTCCAGCTGTACACTCAAGGGCCACAGAGCGGAGCCCTCAAGACC 808
QY 721 ACAGTCATCTCATCTAGCTTTCTTTGGCTGCTGCTGCCATATTATGTTGGGATCAGC 780
Db 809 ACAGTCATCTCATCTAGCTTTCTTTGGCTGCTGCTGCCATATTATGTTGGGATCAGC 868
QY 781 ATGAGCTCTCTCATCTTTTGGGAGTCATCAGAGGATGTGACTTCGAGAGCATTTGG 840
Db 869 ATGAGCTCTCTCATCTTTTGGGAGTCATCAGAGGATGTGACTTCGAGAGCATTTGG 928
QY 841 CACAAGTGGATCTCCATCAGAGGCGCTGCTCTTCCACTGTTGCTGAGCCCATC 900
Db 929 CACAAGTGGATTTCCATCAGAGGCGCTAGCTTTCTTCCACTGTTGCTGAGCCCATC 988
QY 901 CTCTATGCTCTCTCGGGCCAGTTCAAGAGCTCTGCCAGCATCAGCTCACTCATG 960
Db 989 CTCTATGCTTTCTTGGAGCCAAATTAAACCTCTGCCAGCAGCATCTCACTCTG 1048
QY 961 AGCAGAGGCTCCAGCTCAAGATCTTTCCAAAGGAAAGGGGGTGCACACTCTTCCGTC 1020
Db 1049 AGCAGAGGCTCCAGCTCAAGATCTTTCCAAAGGAAAGGGGGTGCACACTCTTCT 1108
QY 1021 TCACAGAGTCAAGATCTCTCAAGTTTCACTCCAGCTAACCCCTATGCAAGACTTAT 1080
Db 1109 TCACAGTCAAGTCTCAAGTTTCACTCCAGCTAA-----CACAGATGTAAAA 1160
QY 1081 AAT 1140
Db 1161 GACTTTTATATACGATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1220
QY 1141 ACTGACAGCTCTGTACAGTTTTTTTTTTTTTTTTTTTTTAATTAATTAATTAATTA 1200
Db 1221 ACTGACCAATATGTACAGTTTTTTTATGCTTGTGTTGATTTTGT-----CTTGTGTT 1274
QY 1201 TCTAGTTTTTGTGAGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
Db 1275 TTTAGTTTTTGTGAGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1330
QY 1261 GAATGAGCGTCTAGGAGGAGCTGTGCGCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
Db 1331 TGATGTGTCTAGGAGGAGCTGTGCGCAAGTTCTTAGTTGTCTGTGTGTGTAG 1390
QY 1321 GACTGTAGACTGTAGAGGAGAAACTGGAACATTCAGAAATGTGTGTAATTAATTAATA 1380
Db 1391 GACTGTAGAA-----AAGGGAACTGAACATTCAGAGCGTGTAGTAATCACTAAA 1442
QY 1381 GCTAGCCGTCATCTCAGCTGTTGCTGCATA 1411
Db 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

RESULT 8
AAI65467
ID AAI65467 standard; DNA; 1679 BP.
XX AC AAI65467;
XX AC AAI65467;
DT 10-DEC-2001 (first entry)
DE Nucleotide sequence of a human polynucleotide.
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;
KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.
OS Homo sapiens.
PN WO200164752-A2.
XX 07-SEP-2001.
FD

XX 28-FEB-2001; 2001WO-US006322.
XX 02-MAR-2000; 2000US-00517605.
XX (UUNY) UNIV NEW YORK STATE.
XX (UUNI-) UNIV NIJMEGEN.
XX Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX WPI; 2001-602565/68.
XX An antibody for the treatment or prevention of HIV-infection comprises a
XX gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of
XX DC-SIGN due to concomitant conformational change.
XX Disclosure; Page 126-127; 131pp; English.
XX The specification describes an antibody which is specific for an
XX antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or
XX is exposed upon gp120 binding of DC-SIGN due to concomitant
XX conformational change. DC-SIGN is a receptor that is specifically
XX expressed on dendritic cells and facilitates infection of T lymphocytes
XX with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC
XX -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)
XX with high affinity. The antibody of the invention inhibits the trans
XX enhancement of HIV entry into a T cell or macrophage facilitated by
XX dendritic cells. The antibody is useful to treat or prevent HIV
XX infection. The present sequence represents a human polynucleotide, which
XX is used in the course of the invention
XX SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 55.1%; Score 933.4; DB 4; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTCTGTGATACTCTGGAAGAGTGGGCTGAGAGACTATGACTCCAAACAAG 60
Db 104 ATATACACTCTGTGATACTCTGGAAGAGTGGGCTGAGAGACTATGACTCCATGAG 163
QY 61 GAACCTCTCTTCCGGATGAAACGTCATTTCAATAGGATCTTCCTGCCCACTATCTAC 120
Db 164 GAACCTCTCTTCCGGATGAAACGTCATTTCAATAGGATCTTCCTGCCCACTATCTAC 223
QY 121 TTCTATCTCTTCTGACTGCGCATAGTCGGCAATGGATTGGTGTGATCCTGCTGATGGTTAC 180
Db 224 TCCATCATCTCTTCTTAACTGGCATTTGGGGCAATGGATTGGTGTGATCCTGCTGATGGTTAC 283
QY 181 CAGAGAAAGCTAAGGAGCATGACGCAAGTACCGGCTGACCTGTCTAGTGGCTGACCTC 240
Db 284 CAGAGAAAGCTAAGGAGCATGACGCAAGTACCGGCTGACCTGTCTAGTGGCTGACCTC 343
QY 241 CTCCTTTGTCATCAGCTCCCTCTTGGGCAATGGATTGGTGTGATCCTGCTGATGGTTAC 300
Db 344 CTCCTTTGTCATCAGCTCCCTCTTGGGCAATGGATTGGTGTGATCCTGCTGATGGTTAC 403
QY 301 AAATTTTGTGTGAGGCTGTCCATATCATCTACATCTGTAACCTCTACAGAGCGTTCTC 360
Db 404 AACTTCTTATGCAAGGAGTCCATGTCTATACAGATCAGCTCAGCTGATGGTTCTC 463
QY 361 ATCTTGGCTTATCAGCTGCAAGGAGTCCATGTGCTGAGTGGGCTGATGGTGTGATGG 420
Db 464 ATCTTGGCTTATCAGCTGCAAGGAGTCCATGTGCTGAGTGGGCTGATGGTGTGATGG 523
QY 421 AGGCCAAGGAAAGTCTGCTGCAAGGAGTCCATGTGCTGAGTGGGCTGATGGTGTGATGG 480
Db 524 AGGCCAAGGAAAGTCTGCTGCAAGGAGTCCATGTGCTGAGTGGGCTGATGGTGTGATGG 583
QY 481 CTCCTGACTATACCTGACTTTCATCTTTTCCGACCTCAGCCAGGGGAGACATCAGTCAGGG 540
Db 584 CTGCTGACTATTCCTGACTTTCATCTTTTCCGACCTCAGCCAGGGGAGACATCAGTCAGGG 628

QY 541 GATGACAGGTACATCTGTGACCGGCTTTATCCCGATAGCCTGTGGATGGTGGTGTTCAA 600
Db 629 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGGTGTGTTCCAG 688
QY 601 TTCAGCATATATGTTGGGTCTCATCTGCGCGGCAATCTATCTCTCTCTCTCTCTCTCT 660
Db 689 TTTGAGACATCATGTTGGCTTATCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 748
QY 661 ATCATCATCTTAAGCTGTACATCTCAAGGCGCCACAGAGCGCAAGGCGCTTCAAGAG 720
Db 749 ATTATCATCTCAAGCTGTACATCTCAAGGCGCCACAGAGCGCAAGGCGCTTCAAGAG 808
QY 721 ACAGTATCTCTATCTGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 809 ACAGTATCTCTATCTGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 868
QY 781 ATCGACTCTCTATCTGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 869 ATCGACTCTCTATCTGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
QY 841 CACAAGTGGATCTCCATCAGAGGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 929 CACAAGTGGATCTCCATCAGAGGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
QY 901 CTCTATGCTTCTCTGCGGCGCAAGTTCAAAAGCTCTGCGGAGCATGCACTCACTCATG 960
Db 989 CTCCTATGCTTCTCTGCGGCGCAAGTTCAAAAGCTCTGCGGAGCATGCACTCACTCATG 1048
QY 961 AGCAGAGGCTCCAGCTCAAGATCTTCTTCAAGGAGGCGGCTGCACTCTTCCGCTC 1020
Db 1049 AGCAGAGGCTCCAGCTCAAGATCTTCTTCAAGGAGGCGGCTGCACTCTTCCGCTC 1108
QY 1021 TCCAGGAGTCAGATCTCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 1080
Db 1109 TCCAGTCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTCCTGAGTTC 1160
QY 1081 AAT 1140
Db 1161 GACTTTTTTTTATACGAT 1220
QY 1141 ACTGACGAGTCTGTACAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
Db 1221 ACTGACCAATATGTACAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1274
QY 1201 TCTAGTTTTTGTAGGTTTCTGACTTATATATATATATATATATATATATATATATAT 1260
Db 1275 TTAGTTTTTGTAGGTTTCTGACTTATATATATATATATATATATATATATATATAT 1330
QY 1261 GAATGAGCGTCTAGGAGGACCTGTGCGCAAGTCTTCTAGTGTGTTTCTGTGTGTAG 1320
Db 1331 TGAATGTGTCTAGGAGGACCTGTGCGCAAGTCTTCTAGTGTGTTTCTGTGTGTAG 1390
QY 1321 GACTGTAGAACTGTAGAGGAGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 1380
Db 1391 GACTGTAGAA-----AAGGGAAGTGAACATTCAGAGCGGTGTAGTGAATCAAGTAA 1442
QY 1381 GCTAGCGGTGATCTCTGAGTGTGCTGCATA 1411
Db 1443 GCTAGAAATGATCCCGAGCTGTTATGCTA 1473

RESULT 9
ABS53992
ID ABS53992 standard; cDNA; 1679 BP.
XX ABS53992;
AC ABS53992;
XX ABS53992;
DT 02-DEC-2002 (first entry)
XX
DE DNA encoding human CXC chemokine receptor 4 (CXCR4).
XX CXC chemokine receptor 4; CXCR4; chemotaxis induction; SDF1a;
KW stromal cell-derived factor 1alpha; angiogenesis; atherosclerosis;
KW

KW restenosis; ischaemic stroke; spinal cord injury; infection; ulcer;
KW human immunodeficiency virus; HIV; acquired immunodeficiency syndrome;
KW AIDS; pain; cancer; benign prostatic hypertrophy; diabetes; obesity;
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypertension; urinary retention; osteoporosis; stroke;
KW anginal pectoris; myocardial infarction; benign prostatic hypertrophy;
KW migraine; vomiting; psychomotor disorder; neurological disorder; anxiety;
KW schizophrenia; dyskinesia; Huntington's disease; restenosis;
KW inflammatory disease; rheumatoid arthritis; diabetic retinopathy;
KW inflammatory bowel disease; atherosclerosis; Alzheimer's disease;
KW congestive heart failure; cardiac remodeling; angiogenic diseases;
KW solid tumour; Kaposi Sarcoma; human; gene; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 89..1147
FT /*tag= a
FT /product= "CXCR4"
FT /note= "CXC chemokine receptor 4"
XX
XX US2002107195-A1.
XX
XX 08-AUG-2002.
XX
XX 17-SEP-2001; 2001US-00953692.
XX
XX 21-JUL-1998; 98US-00933596P.
XX 21-JUL-1999; 99US-00358624.
XX (SMIK) SMITHLINE BEECHAM CORP.
XX Gupta SK;
XX
XX WPI; 2002-697879/75.
XX P-PSDB; ABG32977.
XX
XX Inducing chemotaxis of endothelial cells, useful for regulating
XX angiogenesis e.g. for treating cancer, comprises treatment with stromal
XX cell-derived factor 1 alpha.
XX
XX Disclosure; Fig 1; 26pp; English.
XX
XX The invention describes a method of inducing chemotaxis of endothelial
XX cells by treatment with stromal cell-derived factor 1alpha (SDF1a). The
XX method is used for stimulating EC chemotaxis, and thus angiogenesis, and
XX is used for treating atherosclerosis, restenosis, ischaemic stroke and
XX spinal cord injury. Inhibition of this process is useful in treatment and
XX prevention of a very wide range of diseases, such as, infection (by
XX bacteria, fungi, protozoa or viruses such as human immunodeficiency virus
XX (HIV) and acquired immunodeficiency syndrome (AIDS)), pain, cancer and
XX benign prostatic hypertrophy, diabetes, obesity, anorexia, bulimia,
XX asthma, Parkinson's disease, acute heart failure, hypotension,
XX hypertension, urinary retention, osteoporosis, anginal pectoris,
XX myocardial infarction, stroke, ulcers, benign prostatic hypertrophy,
XX migraine, vomiting, psychotic and neurological disorders (e.g. anxiety,
XX schizophrenia) and dyskinesias (e.g. Huntington's disease), inflammatory
XX diseases, rheumatoid arthritis, diabetic retinopathy, inflammatory bowel
XX disease, atherosclerosis, restenosis, Alzheimer's disease, congestive
XX heart failure, cardiac remodeling, angiogenic diseases, solid tumours,
XX and Kaposi Sarcoma. This sequence encodes the human CXC chemokine
XX receptor 4 (CXCR4) polynucleotide
XX
XX Sequence 1679 BP; 407 A; 368 G; 504 T; 0 U; 0 Other;

Query Match 55.1%; Score 933.4; DB 6; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAG 60
Db 104 ATATACACTTCTGATACTACTACCGAGGAATGGCTCAGGGAGTATGACTCCATGAAG 163

Qy	61	GAACCCCTGCTTCCGGATGAAAACGTCATTTCAATAGGATCTTCTGCCCAACCATCTAC	120
Db	164	GAACCCCTGTTTCCGTGAGAAAAATGCTAAATTTCAATAAAATCTTCTCTGCCCAACCATCTAC	223
Qy	121	TTCAATCATCTCTTGTACTGGCATAGTCGGCAATGATTTGGTGTATCTCTGGTCACTAGGGTTAC	180
Db	224	TCCATCATCTTCTTAACTGGCAATTTGGGCANUGAATTTGGTCATCTCTGGTCACTAGGGTTAC	283
Qy	181	CAGAAGAAGCTTAAGGACATGACGGACAAGTACCGGTGCACTCTGTCACTGGCTGACCTC	240
Db	284	CAGAAGAAAATTGAGAAGCATGACGGACAAGTACAGGCTGCACTGTCACTGGCGGACCTC	343
Qy	241	CTCTTTTGTCAATCACATCCCTTCTGGCGAGTTGATGCCATGGCTGACTGGTACTTTGGG	300
Db	344	CTCTTTTGTCAATCACGCTTCCCTTCTGGCGAGTTGATGCCGTGGCAAACTGCTACTTTGGG	403
Qy	301	AAATTTTGTGAAGGTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC	360
Db	404	AACTTCTATGCAAGGCAGTCCATGTCTATCTACAGTCAACCTCTACAGCAGTGTCTC	463
Qy	361	ATCTGGCTTCATCAGCCTGGACCGGTACTCTGCCATTTGCCAGCCACCAACAGTCAA	420
Db	464	ATCTGGCTTCATCAGTCTGGACCGGTACTCTGCCATTCGTCCACGCCACCAACAGTCA	523
Qy	421	AGSCCAAGGAAACTGCTGGGTGAAAAGGCAGTCTATGTGGCGCTCTGGATCCCAAGCCCTC	480
Db	524	AGSCCAAGGAAGCTGTTGGCTGAAAAGTGCTATGTGGCGCTCTGGATCCCTGCCCTC	583
Qy	481	CTCCTGACTATACCTGACTTCATCTTTTCCGACGTTCAGCCAGGGGGACATCAGTCAGGGG	540
Db	584	CTGCTGACTATATCCGACTTCATCTTTTCCAAAG-----TCAGTGAGGCA	628
Qy	541	GATCACAGGTACATCTGTGACCGGCTTTACCCCGATAGCTGTGGATGGTGGTGTTCAA	600
Db	629	GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGGTGTGTCCAG	688
Qy	601	TTCCAGCATATAATGTGGGTCTCATCCTGCCCGGCATCGTCTACTCTCTCTCTCTCTCTCT	660
Db	689	TTTCAGCACATCATGGTTGGCGCTTATCCTGCCTGGTATTGTCTCATCTCTCTCTCTCTCT	748
Qy	661	ATCATCATCTCTAAGCTGTCACTCAAGGGCCACAGAAAGCGACGAGGCCCTCAAGACG	720
Db	749	ATTATCATCTCCAAGCTGTCACTCCAGGGCCACAGAAAGCGACGAGGCCCTCAAGACC	808
Qy	721	ACAGTCACTCTCATCTTAGCTTTCTTTTGCCTGTGCTGGTGCATATTATGTGGGGATCAGC	780
Db	809	ACAGTCACTCTCATCTTAGCTTTCTTTTGCCTGTGCTGGTGCCTTACTACATTGGGATCAGC	868
Qy	781	ATCGACTCTCTCATCTTTTGGAGTTCATCAAGCAGGATGTGACTTCGAGAGCATTTGTG	840
Db	869	ATCGACTCTCTCATCTCTCTGGAATCATCAAGCAGGAGTGTGAGTTTGAGAACACTGTG	928
Qy	841	CACAAGTGGATCTCCATCACAGAGGCCCTCGCCCTTCTTCCACTGTGCTCTGAAACCCCATC	900
Db	929	CACAAGTGGATTTCCATCACAGAGGCCCTTAGCTTTCTTCCACTGTGCTGTAACCCCATC	988
Qy	901	CTCTATGCCCTTCTCGGGGCAAGTTCAAAAGCTCTGCCAGANGCAGTCAACTCATG	960
Db	989	CTCTATGCTTTCTCTTGAGGCAAAATTTAAACCTCTGCCAGCAGCAGTCACTCTGTG	1048
Qy	961	AGCAGAGGCTCCAGCCTCAAGATCCTTTTCAAAAGGAAAGCGGGTGGACACTCTTCGCTC	1020
Db	1049	AGCAGAGGCTCCAGCCTCAAGATCCTCTCCAAAGGAAGCGAGGTGACATTCATCTGTT	1108
Qy	1021	TCCAGGGAGTCAGAACTCCTCCAGTTTTCATCTCCAGCTTAACCCCTTATGCAAGACTTAT	1080
Db	1109	TCCACTGAGTCTGAGTCTTCAAGTTTTCATCTCCAGCTAA-----CACAGATGTAAAA	1160
Qy	1081	AATATATATATATATGATAAGAAACTTTTTTATGTTCACATTTTCCAGATATAGAG	1140
Db	1161	GACTTTTTTTATACGATAAATAACTTTTTTTTTTAAGTTTACACATTTTTCAGATATAAAG	1220
Qy	1141	ACTGACAGCTCTGTACAGTTTTTTTTTTTTTTTTTTTTTAAATGACTCTGGGAGTTATGTTC	1200

DB	1221	ACTGACCAATATTGTACAGTTTTTATTGCTGTGTGGAATTTTGT-----CTGTGTTTC	1274
QY	1201	TCTAGTTTTTGTGAGGTTTGACCTTAATTTATATAAATATTGTTTTTTGTTGTTTTCATGT	1260
DB	1275	TTTAGTTTTTGTGNAGTTTAATTCACCTATTATA---TAAATTTTTTTGTTTTCATAT	1330
QY	1261	GAATGACGCTCTAGGACGACCTGTGGCCAAAGTTCTTAGTAGCTGTTATCTGTGTGTAG	1320
DB	1331	TGATGTGTGCTAGGCAGGACCTGTGGCCAAAGTTCTTAGTTGCTGTATGTCTCTGTGTAG	1390
QY	1321	GACTGTAGAAGCTGTAGAGGAGAAAGCTGACACATTCAGACATGTGCTAAATTTGAATAAA	1380
DB	1391	GACTGTAGAA-----AAGGGAACCTGAACATTCAGAGCGGTAGTGAATCACGTAAA	1442
QY	1381	GCTAGCCGTGATCCTCAGCTGTTGCTGCATA	1411
DB	1443	GCTAGAAATGATCCCGAGCTGTTTATGCATA	1473
RESULT 10			
ABV78039			
XX	ID	ABV78039 standard; DNA; 1679 BP.	
XX	AC	ABV78039;	
XX	AC		
DT	12-NOV-2002	(first entry)	
XX	XX		
DE	XX	Hypoxia-regulated protein coding sequence #59.	
KW	XX	Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;	
KW	KW	antiinflammatory; vulnerrary; gynecological; ophthalmological; vaccine;	
KW	KW	hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;	
KW	KW	ischaemic condition; reperfusion injury; retinopathy; neonatal stress;	
KW	KW	preclapsemia; atherosclerosis; inflammatory condition; wound healing;	
XX	XX	inflammation; erythropoiesis; hair loss; human; gene; ds.	
OS	XX	Homo sapiens.	
XX	XX		
FN	WO200246465-A2.		
XX	XX		
PD	13-JUN-2002.		
XX	XX		
PF	10-DEC-2001; 2001WO-GB005458.		
XX	XX		
PR	08-DEC-2000; 2000GB-00030076.		
PR	08-FEB-2001; 2001GB-00003156.		
PR	25-OCT-2001; 2001GB-00025666.		
XX	XX		
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.		
XX	XX		
PI	White J, Mundy CR, Ward NR, Krize D, Kingsman SM, Harris RA;		
PI	Rayner WN;		
XX	XX		
DR	WPI; 2002-627238/67.		
XX	XX		
PT	Identifying a gene involved in disease for treating hypoxia-regulated		
PT	conditions, comprises comparing the transcriptome/proteome of two cell		
PT	types under different conditions and identifying a differentially		
PT	regulated gene.		
XX	XX		
PS	Claim 37; Page 378; 538pp; English.		
XX	XX		
CC	The present invention relates to methods for identifying genes and		
CC	proteins that are implicated in a specific disease or physiological		
CC	condition. The method comprises comparing the transcriptome/proteome of a		
CC	specialised cell type implicated in a disease or condition with that of a		
CC	second specialised cell type, under two experimental conditions, and		
CC	identifying a gene that is differentially regulated in the two		
CC	specialised cell types under experimental conditions. ABV77873-ABV78116		
CC	and ABP65061-ABP65257 were identified using the methods of the invention.		
CC	The coding sequences and proteins are useful for treating a disease in a		
CC	patient, for manufacture of a medicament for treating hypoxia-regulated		

CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
 CC biological response to hypoxia conditions, or hypoxic-associated
 CC pathology in a patient. The coding sequences and proteins are also useful
 CC for monitoring the therapeutic treatment of a disease or physiological
 CC condition, such as cancer, ischaemic conditions, reperfusion injury,
 CC retinopathy, neonatal stress, pre-eclampsia, atherosclerosis, inflammatory
 CC conditions, wound healing, inflammation, erythropoiesis or hair loss
 XX
 SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match 55.1%; Score 933.4; DB 6; Length 1679;
 Best Local Similarity 81.8%; Pred. No. 4.3e-207;
 Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGTAATACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAAACAAG 60
 DB 104 ATATACACTTCTGTAATACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAAACAAG 163
 QY 61 GAACCCCTGCTCCGGGATGAAGACGTCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 120
 DB 164 GAACCCCTGCTCCGGGATGAAGACGTCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 223
 QY 121 TTATCATCTCTTGTAGTCGATAGTCGGCAATGGATGGTATGATCTCTGTCATGGGTTAC 180
 DB 224 TCCATCATCTCTTAACTGGCATTTGGGCAATGGATGGTATGATCTCTGTCATGGGTTAC 283
 QY 181 CAGAAGAAGCTAAGGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGGTGACCTC 240
 DB 284 CAGNAGAACTGAGAGCATGACGACAGTACCGGCTGCACCTGTCAGTGGGTGACCTC 343
 QY 241 CTCCTTTGTCATCACTCCCTCTCTGGCAGTTGATGCCAATGGCTGACTGGTACTTTGGG 300
 DB 344 CTCCTTTGTCATCACTCCCTCTCTGGCAGTTGATGCCAATGGCTGACTGGTACTTTGGG 403
 QY 301 AATATTTTGTGAAGGCTGCCATATCATCTACATGTCACCTCTACAGCAGCGTTCTC 360
 DB 404 AACTTCTATGCAAGGAGTCCATGTCATCTACATGTCACCTCTACAGCAGCGTTCTC 463
 QY 361 ATCTGCGCTTTCATAGCCTTGGACCGGTACTCTGCCATTTGTCCACGCCCAACAGCTCAA 420
 DB 464 ATCTGCGCTTTCATAGCCTTGGACCGGTACTCTGCCATTTGTCCACGCCCAACAGCTCAA 523
 QY 421 AGCCAGGAACTGCTGGCTGAAAGGAGTCTATGTGGCGTCTGGATCCAGCCCTC 480
 DB 524 AGCCAGGAACTGCTGGCTGAAAGGAGTCTATGTGGCGTCTGGATCCAGCCCTC 583
 QY 481 CTCTGACTATACCTGACTTTCATCTTTGCCAGCTCAGCCAGGGGACATCAGTCAGGGG 540
 DB 584 CTGCTGACTATTCGCACTTTCATCTTTGCCAGCTCAGCCAGGGGACATCAGTCAGGGG 628
 QY 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTTTCAA 600
 DB 629 GATGACAGATATATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTTTCAA 688
 QY 601 TTCCAGCATATATGTTGGGTCTCATCTCCCGCGCATGTCATCTCTCTGTTACTGC 660
 DB 689 TTTGAGCATATATGTTGGGTCTCATCTCCCGCGCATGTCATCTCTCTGTTACTGC 748
 QY 661 ATCATCATCTTAAGTGTGCACACTCCAGGGCCACAGAGCGCAAGGCCCTCAAGAGC 720
 DB 749 ATTATCATCTCAAGCTGTCCACTCCAGGGCCACAGAGCGCAAGGCCCTCAAGAGC 808
 QY 721 ACAGTCATCTCATCTCTGAGTCTTCTTGGCTGCTGCTGCCATATTATGCTGGGATCAGC 780
 DB 809 ACAGTCATCTCATCTCTGAGTCTTCTTGGCTGCTGCTGCCATATTATGCTGGGATCAGC 868
 QY 781 ATGACTCTCTTCTCTTTTGGGAGTGCATCAAGCAAGAGTGTGACTTCGAGAGCATTTGTG 840
 DB 869 ATGACTCTCTTCTCTTCTCTGGAATCATCAAGCAAGGAGTGTGAGTTTGAGAACATGTG 928
 QY 841 CACAGTGGATCTCCATCAAGAGGCGCTCGCTTCTTCCATGCTGGTGCCTGAACCCCATC 900
 DB 929 CACAGTGGATCTCCATCAAGAGGCGCTCGCTTCTTCCATGCTGGTGCCTGAACCCCATC 988

QY 901 CTCTATGCTTCTCTGGGGCCAAAGTTCAAAGCTCTCCAGCATGCACTCAATCCCATG 960
 DB 989 CTCTATGCTTCTCTGGGGCCAAAGTTCAAAGCTCTCCAGCATGCACTCAATCCCATG 1048
 QY 961 AGCAGAGCTCCAGCTCAAGATCTTTCCAAAGGAAGGGGTGGACACTCTCTCCGTC 1020
 DB 1049 AGCAGAGGCTCCAGCTCAAGATCTTTCCAAAGGAAGGGGTGGACACTCTCTCTGTT 1108
 QY 1021 TCCAGGAGTCCAGATCTCCAGTCTTCTCACTCCAGCTAAACCTTTATGCAAGACTTATAT 1080
 DB 1109 TCCAGTCTGAGTCTGAGTCTTCAAGTCTTCTCACTCCAGCTAAACCTTTATGCAAGACTTATAT 1160
 QY 1081 AATATATATATATATGATAAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG 1140
 DB 1161 GACTTTTTTATACGATAAATAACTTTTTTTAAGTTACACATTTTTCAGATATAAAG 1220
 QY 1141 ACTGACAGTCTTGTACAGTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCC 1200
 DB 1221 ACTGACCAATATTTGACAGTTTTTATTGCTGTTGGATTTTTGT-CTTGTGTTTC 1274
 QY 1201 TCTAGTTTTGTGAGGTTTGAATTAATATATAAATATGTTTTTTTGTGTTTTCATGT 1260
 DB 1275 TTTAGTTTTGTGAGGTTTGAATTAATGACTTATTTATA-----TAAATTTTTTTTTCATAT 1330
 QY 1261 GAATGAGCGTCTAGGCGAGGACCTGTGGCCAAAGTTCTTTAGTAGCTGTTTATCTGTGTAG 1320
 DB 1331 TGATGTGTCTAGGCGAGGACCTGTGGCCAAAGTTCTTTAGTTGCTGTATGCTCGGTAG 1390
 QY 1321 GACTGTGAACCTGTAGAGGAAGAACTGAACATTTCCAGAAATGTGTGTAATTAATAA 1380
 DB 1391 GACTGTGAAC-----AAGGGAACCTGAACATTTCCAGAGCGTGTAGTGAATCAACGTAA 1442
 QY 1381 GCTAGCGCTGATCTCAGCTGTTCTGTCATA 1411
 DB 1443 CTTAATAATGATCCCAAGCTGTTTATGATA 1473

RESULT 11

ABX08779

ID ABX08779 standard; cDNA; 1679 BP.

XX AC ABX08779;

XX DT 21-JAN-2003 (first entry)

XX DE Angiogenesis-associated human polynucleotide sequence #41.

XX KW Human; angiogenesis-associated transcript; angiogenesis;

XX KW angiogenesis-associated disease; cancer; cyrostatic; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200279492-A2.

XX PD 10-OCT-2002.

XX PF 14-FEB-2002; 2002WO-US004915.

XX PR 14-FEB-2001; 2001US-00784356.

XX PR 22-FEB-2001; 2001US-00791390.

XX PR 19-APR-2001; 2001US-0285475P.

XX PR 03-AUG-2001; 2001US-0310025P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334244P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Murray R, Glynn R, Watson SR, Aziz N;

XX WPI; 2003-040681/03.

DR P-PSDB; AB03495.

XX Detecting angiogenesis-associated transcript in a cell for diagnosing and
 PT treating cancer by contacting a sample with a polynucleotide that
 PT exhibits changes in expression level as a function of time in tissue
 PT undergoing angiogenesis.

XX Example 2; Page 217-218; 291pp; English.

XX The present invention relates to methods and compositions for detecting
 CC an angiogenesis-associated transcript in a cell in a patient. The method
 CC involves contacting a biological sample from the patient with a
 CC polynucleotide that selectively hybridizes to a sequence at least 80%
 CC identical to any of the angiogenesis-associated human polynucleotide
 CC sequences given in the specification. These angiogenesis-associated
 CC polynucleotide sequences comprise genes that exhibit changes in
 CC expression levels as a function of time in tissue undergoing
 CC angiogenesis. The method and the polynucleotide sequences of the
 CC invention are useful for diagnosing and treating angiogenesis and
 CC angiogenesis-associated diseases e.g. cancer. The polynucleotide
 CC sequences are also useful in the gene therapy of such disorders. The
 CC angiogenesis-associated proteins encoded by the polynucleotide sequences
 CC are useful as a vaccine for therapeutic and prophylactic immunisation.
 CC ABX08739-ABX08853 represent angiogenesis-associated polynucleotide
 CC sequences

XX Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match 55.1%; Score 933.4; DB 7; Length 1679;
 Best Local Similarity 81.8%; Pred. No. 4.3e-207;
 Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCAAACAG 60
 DB 104 ATATACACTTCCAGTAACACCGAGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163
 QY 61 GRACCTGCTCCGGGATGAAACGTCATTTCAATAGATCTTCTGCCACCACTAC 120
 DB 164 GAACCTGTGTCGGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCACCACTAC 223
 QY 121 TTCAATCATCTTCTGACTGGCATAGTCGGCAATGATTTGGTTCATCTGCTCATGGGTTAC 180
 DB 224 TCCATCATCTTCTTAAGTGGCATTTGGGCAATGATTTGGTTCATCTGCTCATGGGTTAC 283
 QY 181 CAGAAGACTAAGAGCATAGCGCAAGATACCGCTGCACTGTCTAGTGGCTGACCTC 240
 DB 284 CAGAAGAACTGAGAAGCATAGCGCAAGATACCGCTGCACTGTCTAGTGGCTGACCTC 343
 QY 241 CTCTTTGTCTACACACTCCCTTCTGGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTG 300
 DB 344 CTCTTTGTCTACACGCTTCCCTTCTGGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTG 403
 QY 301 AAATTTTGTGAAGCTGTCCATATCATCTACATGTCACCTCTACAGCAGCGTTCTC 360
 DB 404 AACTTCTATGCAAGGCACTCCATGTCTATCTACAGTCAACCTCTACAGCAGTGTCTC 463
 QY 361 ATCTTGGCTTCATCAGCTGACCGGTACCTGCGCATTTGTCAGCGCCACCAAGTCAA 420
 DB 464 ATCTTGGCTTCATCAGCTGACCGGTACCTGCGCATTTGTCAGCGCCACCAAGTCAA 523
 QY 421 AGCCCAAGGAACTGTGGCTGAAAGGCACTATGTGGGCGTCTGGATCCCGCCCTC 480
 DB 524 AGCCCAAGGAACTGTGGCTGAAAGGCGTCTATGTGGGCGTCTGGATCCCGCCCTC 583
 QY 481 CTCTCTGACTATACCTGACTTCACTTTTGGCGAGTTCAGCGGGGCAATCATGTCAGGG 540
 DB 584 CTGCTGACTATTTCCGACTTCATCTTTTGGCAAG-----TCAGTGAAGCA 628
 QY 541 GATGACAGGTACATCTGTGACCGCTTTTACCCCGATAGCTGTGATGGTGGTGTTCAA 600
 DB 629 GATGACAGATATATCTGTACCGCTTCTACCCCAATGACTTGTGGTGGTGTGTTCAG 688
 QY 601 TTCCAGCATATTAAGTGGGTCTCATCTGCGCGGCACTGTCATCTCTCTGTACTGTC 660

DB 589 TTTTCAGCACATCATGTTGGCTTATCTCTGCTGGTATTGTCTATCTCTGCTTATTC 748
 QY 661 ATCATCATCTTAAGCTGTCACTCCCAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACG 720
 DB 749 ATTATCATCTCAAGCTGTCACTCCCAAGGGCCACCAAGAGCGCAAGGCCCTCAAGACC 808
 QY 721 ACAGTCATCTCATCTCTAGCTTTCTTTGCTGCTGCTGCTGCATATATATGTTGGGATCAG 780
 DB 809 ACAGTCATCTCATCTCTGCTTTCTTCTGCTGCTGCTGCTGCTTACTACATTTGGATCAG 868
 QY 781 ATCGACTCTTCTCATCTCTTTTGGAGTTCATCAAGCAAGATGTGACTTTTCGAGACATTTG 840
 DB 869 ATCGACTCTTCTCATCTCTCTCTGGAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 928
 QY 841 CACAAGTGGATCTCCATCAACAGAGGCTCTGCTTCTTCCACTGTTGCTTCCCTGAACCCATC 900
 DB 929 CACAAGTGGATTTCCATCAACAGAGGCTCTGCTTCTTCCACTGTTGCTTCCCTGAACCCATC 988
 QY 901 CTCTATGCTTCTCTCGGGGCAAGTTCAAAAGTCTGTGCCAGCATGCACTCAACTCCATG 960
 DB 989 CTCTATGCTTCTCTGAGCCAAATTTAAAACCTCTGTCCAGCAGCAGCACTCACCTCTGTG 1048
 QY 961 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGAAAGCGGGGTGACACTCTTCCGTC 1020
 DB 1049 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGAAAGCGGGGTGACACTCTTCTGTT 1108
 QY 1021 TCCACGAGTCAGAACTCTCCAGTTTCTCACTCCAGCTAACCTTATGCAAGACTTATAT 1080
 DB 1109 TCCACTGAGTCTGAGTCTTCAAGTTTCTCACTCCAGCTAA-----CACAGATGTAATA 1160
 QY 1081 AATATATATATATATATGATAAGAACTTTTATTTATTTATTTATTTATTTATTTATTTAT 1140
 DB 1161 GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTACACATTTTTCAGATATAAAG 1220
 QY 1141 ACTGACGAGCTGTGACAGTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCC 1200
 DB 1221 ACTGACCAATATTTGACAGTTTTTTTATTTGCTTTGTTGTTGTTTGT-----CTTGTGTT 1274
 QY 1201 TCTAGTTTTGTGAGGTTTGAATTAATAATAATAATAATAATAATAATAATAATAATAATA 1260
 DB 1275 TTTAGTTTTGTGAAGTTTAAATGACTTATTATA-----TAAATTTTTTTTGTTCATAT 1330
 QY 1261 GAATGAGCTGTAGGAGGAGCTGTGGCCAAAGTTCTTAGTAGCTTTATCTGTTGTTAG 1320
 DB 1331 TGATGTGTGTGTAGGAGGAGCTGTGGCCAAAGTTCTTAGTTGCTGTATGTCTCGTGTAG 1390
 QY 1321 GACTGTAGAACTGTAGAGGAGAAACTGAACATTTCCAGAAATGTGTGTAATAATGAATAA 1380
 DB 1391 GACTGTAGAA-----AAGGGAACTGAACATTTCCAGAGCGTGTAGTGAATCACGTAAA 1442
 QY 1381 GCTAGCGCTGATCTCTAGCTGTGCTGCATA 1411
 DB 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

RESULT 12
 ABX74454

ID ABX74454 standard; cDNA; 1679 BP.

XX AC ABX74454;

XX DT 21-MAR-2003 (first entry)

XX Human cDNA sequence #26 up-regulated in CC-RCC patients.

XX Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.

OS Homo sapiens.

XX W0200279411-A2.

```

XX PD 10-OCT-2002.
XX PF 29-MAR-2002; 2002WO-US009576.
XX PR 29-MAR-2001; 2001US-0279411P.
XX PA (VAND-) VAN ANDEL INST.
XX PI Haab B, Rhodes D, Teh BT, Takashi M;
XX DR WPI; 2003-040679/03.
XX PT New microarray, comprising a matrix of cDNA probe from a set of probes
XX PT immobilized to a solid surface in predetermined order, useful in the
XX PT prognosis of patients with clear cell renal carcinoma.
XX PS Claim 35; Page 132-133; 179pp; English.
XX CC The present invention relates to a microarray comprising a matrix of at
XX CC least one cDNA probe from a set of probes immobilised to a solid surface
XX CC in a predetermined order, where a row of pixels corresponds to replicates
XX CC of one distinct probe from the set. The probes are complementary to
XX CC nucleic acid sequences that are expressed differentially in aggressive as
XX CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)
XX CC and that hybridise to the probes under high stringency conditions. The
XX CC microarray is useful for the prognosis of patients with CC-RCC, wherein
XX CC aggressive and non-aggressive CC-RCC tumour types are characterised by
XX CC differential expression profiles of genes that hybridise with one or more
XX CC probes immobilised on the microarray. The arrays are useful for gene
XX CC expression profiling of tumour and normal tissues. The present sequence
XX CC represents a human cDNA sequence up-regulated in CC-RCC patients
XX SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match 55.18; Score 933.4; DB 7; Length 1679;
Best Local Similarity 81.8; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGTCTGGAGACTATGACTCCCAACAAG 60
Db 104 ATATACACTTCAGATACTACACCGAGGAATGGCTCAGGGACTATGACTCCATGAAG 163

Qy 61 GAACCCCTCTCCGGATGAAGAGCTCCATTCAATAGGATCTCTCTGCCACCATCTAC 120
Db 164 GAACCCCTGTTCCGGTGAAGAAATGCTAAATTTCAATAAAATCTCTCTGCCACCATCTAC 223

Qy 121 TTCATCATCTTCTGACTGGCATAGTCGCAATGGATTGGTGATCTCTGGTCACTGGGTTAC 180
Db 224 TCCATCATCTTCTTAATCTGGCATTTGGGCAATGGATTGGTCACTCTGGTCACTGGGTTAC 283

Qy 181 CAGAAAGCTTAGGAGCATGACGGACAAGTACCGGCTGCACTGTCACTGGGTGACCTC 240
Db 284 CAGAAAGAACTGAGAAGCATGACGGACAAGTACCGGCTGCACTGTCACTGGGTGACCTC 343

Qy 241 CTCCTTCTCATCACCTCCCTCTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
Db 344 CTCCTTCTCATCAGCTTCTCTCTGGCAGTTGATGCCATGGCTGACTGGTACTTTGGG 403

Qy 301 AAATTTTGTGTAGGCTGTCATATCATCTACATCTGTCACCTCTACAGCAGCTTCTC 360
Db 404 AACTTCTATGCAAGGCAGTCCATGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTC 463

Qy 361 ATCTCGCCCTCATACGCTGGACCGGTACTCTGCCATTTGTCACGCCACCAACAGTCAA 420
Db 464 ATCTCGCCCTCATACGCTGGACCGGTACTCTGCCATTCGTCACGCCACCAACAGTCAAG 523

Qy 421 AGGCCAAGGAACCTGCTGGCTGAAAAGCAGTCTATGTGGGCGTCTGTGATCCAGGCCCTC 480
Db 524 AGGCCAAGGAAGCTGTTGGCTGAAAAGCTGTCTATGTGGGCTGTGATCCCTGCGCTC 583

Qy 481 CTCCTGATATACCTGATCTCACTTTTGGCAGTGTCTAGCCAGGGGACATCAGTAGGGG 540

```

RESULT 13
 ABZ68886
 ID ABZ68886 standard; cDNA; 1679 BP.
 XX
 AC ABZ68886;
 XX
 DT 28-MAY-2003 (first entry)
 XX
 DE Nucleotide sequence of human chemokine receptor CXCR4.
 XX

```

Db 584 CTGCTGACTATTCGCGACTTCACTTTTGCCAAAG-----TCAGTGAGGCA 628
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCCTGTGGATGGTGGTTCAA 600
Db 629 GATGACAGATATATCTGTGACCGCTTTACCCCAATGACTTGTGGTGGTGTTCAG 688
Qy 601 TTCAGCATATATAGTGGGTCTCATCTGCGCCGGCATCGTCATCCTCTCTGTATGTC 660
Db 689 TTTCAGCACATCATGGTTGGCTTTATCTGCTCGTATTTGTCATCTGTCTGTATGTC 748
Qy 661 ATCATCATCTCTAAGCTGTCACTCCAAAGGCCACCAAGAGCGCAAGGCCCTCAAGACG 720
Db 749 ATTATCATCTCCAGCTGTCACTCCAAAGGCCACCAAGAGCGCAAGGCCCTCAAGACC 808
Qy 721 ACAGTCATCTCATCTCTAGCTTTCTTGGCTGTGGCTGCGCATATATATGTGGGATCAGC 780
Db 809 ACAGTCATCTCATCTCTAGCTTTCTTGGCTGTGGCTGCGCATATATATGTGGGATCAGC 868
Qy 781 ATCGACTCTCTCATCTCTTTGGGAGTCAATCAAGCAAGATGTGACTTCGAGAGCATTTGTG 840
Db 869 ATCGACTCTCTCATCTCTCTCTGGAATCATCAAGCAAGGTGTGAGTTGAGAACACTGTG 928
Qy 841 CACAAGTGGATCTCCATCAGAGAGCGCTGCTTCTTCCACTGTTCGCTGAAACCCCATC 900
Db 923 CACAAGTGGATTTCCATCAGAGAGCGCTGCTTCTTCCACTGTTCGCTGAAACCCCATC 988
Qy 901 CTCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCACTCAATCCATG 960
Db 989 CTCTATGCTTCTCTCGGGGCCAAGTTCAAAAGCTCTGCCAGCATGCACTCAATCCATG 1048
Qy 961 AGCAGAGCTCCAGCTCAAGATCTTTCCAAAGGAAGGGGTGGACACTCTTCGGTC 1020
Db 1049 AGCAGAGGGTCCAGCTCAAGATCTTTCCAAAGGAAGGGGTGGACACTCTTCATCTGTT 1108
Qy 1021 TCCAGGAGTCAAGATCTTCCAGTTTTCACCTCCAGCTTAACCTTATGCAAGACTTATAT 1080
Db 1109 TCCAGTGTGCTGAGTCTTCAAGTTTTCACCTCCAGCTAA-----CACAGATGTAAA 1160
Qy 1081 AATATATATATATATATGATAAAGAACTTTTATGTTTATGTTTCCAGATATAAGAG 1140
Db 1161 GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTACACATTTTTCAGATATAAAG 1220
Qy 1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTC 1200
Db 1221 ACTGACCAATATTTACAGTTTTTTTATGCTGTGTGGATTTTGT-----CTTGTGTTT 1274
Qy 1201 TCTAGTTTTTGTGAGGTTTGACTTAAATTTATATAAATATTGTTTTTTTGTGTTTCATGT 1260
Db 1275 TTTAGTTTTTGTGAGTTTAAATGACTTATTATA---TAAATTTTTTTTGTTCATAT 1330
Qy 1261 GAATGAGGCTTAGCAGGACCTGTGGCAAGTTCTTTAGTAGCTGTTTATCTGTGTAG 1320
Db 1331 TGATGTGTCTTAGCAGGACCTGTGGCAAGTTCTTTAGTTGCTGTATGTCTCGGTAG 1390
Qy 1321 GACTGTAGAACTGTAGAGGAAGAACTCAACATTCAGAAATGTGTGTAATTAATAA 1380
Db 1391 GACTGTAGAA-----AGGGAACTGAACATTTCCAGAGCGTGTAGTAATCAGTAAA 1442
Qy 1381 GCTAGCGCTGATCTCAGCTGTTCGCATA 1411
Db 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

```

KW Human; chemokine receptor; CXCR4; viral infection; surface protein;
KW respiratory virus infection; respiratory syncytial virus infection;
KW RSV infection; bronchiolitis; bronchitis; pneumonia; asthma; gene; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 89..1147
FT /*tag= a
FT /product= "CXCR4"
XX
PN W02003014153-A2.
XX
PD 20-FEB-2003.
XX
PF 12-AUG-2002; 2002WO-CA001248.
XX
PR 10-AUG-2001; 2001US-0311088P.
XX
PA (TOPI-) TOPIGEN PHARM INC.
XX
XX Renzi P, Zenzoumi K;
PI WPI; 2003-256541/25.
XX P-PSDB; ABP97733.
DR
DR
XX
XX
PT Modulating viral infection of a cell, for treating or preventing
PT respiratory virus infections, bronchitis, pneumonia or asthma, by
PT modulating a binding interaction between a cell chemokine-receptor and a
PT surface protein of the virus.
XX
XX Disclosure; Page 117-119; 120pp; English.
XX
XX The present sequence encodes human chemokine receptor CXCR4. The
CC specification describes a method for modulating viral infection of a
CC cell. The method comprises modulating a binding interaction between a
CC cell chemokine-receptor and a surface protein of the virus. The proviso
CC is that the cell chemokine-receptor is not CX3CR1 and that the virus is
CC not HIV. The method is useful for treating or preventing respiratory
CC virus infection in vertebrates, more particularly respiratory syncytial
CC virus (RSV) infections, and related diseases, e.g. bronchiolitis,
CC bronchitis, pneumonia or asthma
XX
SQ Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
Query Match 55.1%; Score 933.4; DB 7; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTTCTGATACTACTTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAG 60
DB 104 ATATACACTTCCAGATAAATACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 163
QY 61 GAACCTGTCTCCGGGATGAAGAGTCCATTTCAATAGGATCTTCCTGCCCAACCATCTAC 120
DB 164 GAACCTGTCTCCGGGATGAAGAGTCCATTTCAATAGGATCTTCCTGCCCAACCATCTAC 223
QY 121 TTCATCATCTTTCTGACTGGCATAGTCGGCAATGGATGGTGATCTCGGTGATCGGGTTAC 180
DB 224 TCCATCATCTTTCTGACTGGCATAGTCGGCAATGGATGGTGATCTCGGTGATCGGGTTAC 283
QY 181 CAGAGAGCTAAGGAGCATGACCGACAAAGTACCGGTGACCTGTGACGTGGCTGACCTC 240
DB 284 CAGAGAGAACTGAGAGCATGACCGACAAAGTACCGGTGACCTGTGACGTGGCTGACCTC 343
QY 241 CTCCTTTGTATCATCACTCCCTTCTGGGCAGTTGATGCCATGGCTGACTGTACTTTGGG 300
DB 344 CTCCTTTGTATCATCACTCCCTTCTGGGCAGTTGATGCCATGGCTGACTGTACTTTGGG 403
QY 301 AAAATTTTGTGATAGGCTGTCATATCATCTACACTGTCAACCTCTACAGGAGGTCTC 360
DB 404 AACTTCTATGCAAGGAGTCCATGTCTATCTACAGTCAACCTCTACAGGAGGTCTC 463

QY 361 ATCTGGCTTCATCAGCTGGACGGTACTCGCCATTGTCCACGCCACCAACAGTCAA 420
DB 464 ATCTGGCTTCATCAGCTGGACGGTACTCGCCATTGTCCACGCCACCAACAGTCAA 523
QY 421 AGGCCAAGGAAGTCTGGCTGAAAAGGCAAGTCTATGTGGCGTCTGGATCCCAAGCCCTC 480
DB 524 AGGCCAAGGAAGTCTGGCTGAAAAGGCTGTATGTGGCGTCTGGATCCCTGCCCTC 583
QY 481 CTCCTGACTATACCTGACTTCTCTTTGGCGAGCTGACGAGGGGAGCATCAGTCAAGGG 540
DB 584 CTGTGACTATTCGCGACTTCTCTTTGCCAAGC-----TCAGTGAGGCA 628
QY 541 GATGACGCTACATCTGTGACCGCTTTTACCCCATAGCCTGTGATGGTGGTGTTCAA 600
DB 629 GATGACGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGTGGTGTTCAG 688
QY 601 TTCAGCAATATAATGGTGGGCTCATCTGCGCGAGCTGACGAGGGGAGCATCAGTCAAGGG 660
DB 689 TTTGAGCAGATCATGGTGGGCTTATCTGCTGCTGTATTTGTCATCTGCTGCTGATTCG 748
QY 661 ATCATCATCTCTAAGCTGTGACACTTCCAAAGGGCCACAGAGGCGCAAGGGCCCTCAAGACG 720
DB 749 ATTATCATCTCCAGCTGTCACTTCCAAAGGGCCACAGAGGCGCAAGGGCCCTCAAGACC 808
QY 721 ACAGTCATCTCTCATCTAGCTTTCTTTGCTGGCTGGCTGCTATTAATGTGGGATCAGC 780
DB 809 ACAGTCATCTCTCATCTAGCTTTCTTTGCTGGCTGGCTGCTTACTACTCATTTGGGATCAGC 868
QY 781 ATCGACTCTCTCATCTTTGGGAGTCAATCAAGCAAGGATGTGACTTCGAGGAGATCTG 840
DB 869 ATCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 928
QY 841 CACAAGTGGATCTCCATCAGAGGCGCTCGCTCTTTCCACTGTGTGCTGAAACCCCATC 900
DB 929 CACAAGTGGATTTCCATCAGAGGCGCTCTAGCTTTCTTCCACTGTGTGCTGAAACCCCATC 988
QY 901 CTCTATGCTCTCTCGGGCCAGTTCAAGAGCTCTGCCAGCATGCACTCACTCAATG 960
DB 989 CTCTATGCTCTCTCGGGCCAGTTTAAACCTTGTGCCAGCAGCAGCTCACTCACTCTGTG 1048
QY 961 AGCAGAGCTCCAGCTCAAGATCTTTCCAAAGGAAGCGGGTGGAGACTCTTCCCTC 1020
DB 1049 AGCAGAGGTCAGGCTCAAGATCTCTCCAAAGGAAGCGAGGTGGAGACTCTCATCTGT 1108
QY 1021 TCCAGGAGTCAGAAATCTCCAGTTTTCCTCCAGCTAACCTTATGCAAGAGACTTATAT 1080
DB 1109 TCCAGTGTCTGAGTCTTCAAGTTTTCCTCCAGCTAA-----CACAGATGTAAAA 1160
QY 1081 AATATATATATATATATGATAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG 1140
DB 1161 GACTTTTTTATAGCATAAATTAATCTTTTTTAAAGTTTACATTTTTCAGATATAAAG 1220
QY 1141 ACTGACCACTCTGTACAGTTTTTTTTTTTTTTTTTTTAAATGACTGTGGGAGTTTATGTCC 1200
DB 1221 ACTGACCAATATTTGACAGTTTTTATGCTGTGTGGAATTTTGT-----CTTGTGTTTC 1274
QY 1201 TCTAGTTTTGTAGAGTTTGTACTTAATTAATTAATATGTTTGTGTTTGTGTTTTCATGT 1260
DB 1275 TTTAGTTTTGTGAAGTTTAAATGACTTATTATA-----TAAATTTTTTTTGTTCATAT 1330
QY 1261 GAATGAGCTCTPAGGAGGAGCTGTGGCCAGCTTCTTAGTAGCTGTCTTATCTGTGTAG 1320
DB 1331 TGTATGTGTCTAGGAGGAGCTGTGGCCAGCTTCTTAGTGTCTGTATGTCTCGGTAG 1390
QY 1321 GACTGTAGAACTGTAGAGGAAGAAATGAAATTCAGATGTGTGGTGAATTAATTAATA 1380
DB 1391 GACTGTAGAA-----AAGGGAAGTGAACATTCAGAGCGGTGTAGTGAATCAGGTAAA 1442
QY 1381 GCTAGCGGTATCTCAGCTGTGTGTCATA 1411
DB 1443 GCTAGAAATGATCCCGAGCTGTTTATGATA 1473

RESULT 14

ABZ42642
ID ABZ42642 standard; DNA; 1679 BP.
XX
AC ABZ42642;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human CXCR chemokine receptor 4 nucleotide SEQ ID NO:75.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US050107.
XX
PR 19-DEC-2000; 2000US-0257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burner GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81796.
XX

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure: Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, ulcers, or hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABZ42642 to ABZ42649 encode GPCR proteins given in ABP8175 to ABP82019, which are used in the exemplification of the present invention

Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;

Query Match	55.1%;	Score	933.4;	DB	7;	Length	1679;
Best Local Similarity	81.8%;	Pred. No.	4.3e-207;				
Matches	1154;	Conservative	0;	Mismatches	216;	Indels	41;
Gaps	5;						
QY	1	ATATACACTTCTGATAACTACTCTCAAGAAGTGGGGCTGGAGACTATGACTCCAAACAAG	60				
DB	104	ATATACACTTTCAGATTAACATACACCGAGAAATGGGCTCAGGGAGACTATGACTCCATGAAG	163				
QY	61	GAACCCCTGCTCCGGGATGAAGAGTCCATTTTCATATAGGATCTTCTCTGCCACCATCTTAC	120				
DB	164	GAACCCCTGTTTCCCGTGAAGAAATGCTAAATTTCAATAAAATCTTCTCTGCCACCATCTTAC	223				
QY	121	TTTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTTGGTGTGATCTCTGGTCTATGGGTTC	180				
DB	224	TCCATCATCTTCTTAACTTGGCATTTGTGGCAATGGATTTGGTTCATCTCTGGTCTATGGGTTC	283				
QY	181	CAGAAGAGCTAAGAGCATGACGCAAGTAGTACCGCTGCACCTGTCTAGTGGTGTACCTC	240				
DB	284	CAGAGAAACTCAGAAGCATGACGACAAGTAGTACAGCTGCACCTGTCTAGTGGCGGACCTC	343				
QY	241	CTCTTTGTTCATCACACTCCCTTCTGGGCAGTTGATGTCATGGCTGACTGCTGCTACTTTGGG	300				
DB	344	CTCTTTGTTCATCACGCTTCCCTTCTGGGCAGTTGATGCGGTGGCAAACTGGTACTTTGGG	403				
QY	301	AAATTTTGTGTAAAGCTGTCCATATCATCTACATGTCAACCTCTACAGCAGCGTTCTC	360				
DB	404	AACTTCTTATGAAGCGAGTCCATGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTC	463				
QY	361	ATCTGGGCTTCATCAGCGCTGACCGGTACTCTGGCCATTTGTCCACGCCACCAACAGTCAA	420				
DB	464	ATCTGGGCTTCATCAGTCTGACCGCTACTCTGGCCATGCTCCACGCCACCAACAGTCTAG	523				
QY	421	AGGCCAAGGAACTGCTGGCTGAAAGGAGCTATGTATGGGGGTCTGGATCCAGCGCTC	480				
DB	524	AGGCCAAGGAACTGTTGGCTGAAAGGAGTGTATGTTGGGGTCTGGATCCCTGGCTC	583				
QY	481	CTCTGACTATACCTGACTTCACTTTTGGCGAGCTCAGCCAGGGGAGATCATCTAGCGGG	540				
DB	584	CTGTGACTATTCCCAGCTTCATCTTTGCCAAG-----TCAGTAGGCA	628				
QY	541	GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCCTGTGGATGTGGTGTTCAA	600				
DB	629	GATGACAGATATCTGTGACCGCTTTACCCCAATGACTTTGGGGTGGTTGTTTCCAG	688				
QY	601	TTTCCAGCATATAAGTGGTGGGTCTCATCTCTGCCGCGCATCGTTCCTCTCTGTTTACTGC	660				
DB	689	TTTCAGCACATCATGGTTGGCCTTATCCTGCTGGTATTGTTCATCTCTGTCTGTATTGTC	748				
QY	661	ATCATCTCTTAAGCTGTCACTTCCAGGGGCCACGAGGCGAAGGCCCTTCAAGACG	720				
DB	749	ATTATCATCTCAAGCTGTCACTTCCAGGGGCCACGAGGCGCAAGGCCCTTCAAGACC	808				
QY	721	ACAGTCACTCTCATCTAGCTTTCTTTGGCTGCTGGCTGCCATTTATGTGGGGATCAGC	780				
DB	809	ACAGTCACTCTCATCTGGCTTTCTTGGCCTGTGGCTGCCCTACTACATTTGGGATCAGC	868				
QY	781	ATCGACTCTTTCATCTTTGGGATGTCATCAGCAAGGATGTGACTTCGAGAGCATTTGTG	840				
DB	869	ATCGACTCTTTCATCTCTCTGGAATCATCAGCAAGGATGTGAGTTTGAGAACACTGTG	928				
QY	841	CACAAGTGGATCTCCATCAGAGGCGCTTCTTCTTCCACTGTTGCTGTGAACCCCATC	900				
DB	929	CACAAGTGGATTCCATCACCAGGCGCTTAGCTTTCTTCCACTGTTGTGTGAACCCCATC	988				
QY	901	CTCTATGCTCTCTCGGGCCAGTTCAAAAGCTCTGCCACAGTGCATCTCAACTCCATG	960				
DB	989	CTCTATGCTCTCTTGGAGCCAAATTTAAACCTCTGCCACAGTGCATCTACCTCTGTG	1048				
QY	961	AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAAGCGGGGTGGA CACTTTCCTGTC	1020				
DB	1049	AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAAGCGAGGTGGA CATTTCATCTGT	1108				

QY 1021 TCACGAGTCAGAACTCTCCAGCTTTTCACTCCAGCTAACCCCTTATGCAAGACTTATAT 1080
Db 1109 TCACCTGAGTCTGAGTCTTCAAGTCTTCACTCCAGCTAA-----CACAGTGTAA 1160
QY 1081 AATATATATATATATATGATAAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG 1140
Db 1161 GACTTTTITATACGATAAATAACTTTTTTTTAAAGTTACACATTTTCCAGATATAAAG 1220
QY 1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTTTAAATGACTGTTGGAGTTTATGTTCC 1200
Db 1221 ACTGACCAATATTGTACAGTTTTTATGCTGTGTGGAATTTTGT-----CTGTGTTTC 1274
QY 1201 TCTAGTTTTTGTGAGGTTTGACTTAATTTATATAAATATTGTTTTTGTGTTTCATGT 1260
Db 1275 TTTAGTTTTTGTGAGTTTAAATGACTTATTATA---TAAATTTTTTTTGTGTTTCATAT 1330
QY 1261 GAATGAGCTGTAGGAGGACCTGTGGCCAAAGTTCTTAGTGTGTTTATCTGTGTAG 1320
Db 1331 TGATGTGTGTAGGAGGACCTGTGGCCAAAGTTCTTAGTGTGTTTATCTGTGTAG 1390
QY 1321 GACTGTAGAACTGTAGAGGAGAACTGAACATTCCAGAAATGTGTGTAATTAATAA 1380
Db 1391 GACTGTAGAA-----AAGGAACTGAACATCCAGAGCGGTAGTGTATCACGTAA 1442
QY 1381 GCTAGCCGTGATCTCAGCTTTGCTGCATA 1411
Db 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

RESULT 15
ADC98645
ID ADC98645 standard; cDNA; 1679 BP.
AC ADC98645;
XX 01-JAN-2004 (first entry)
DT Human CXc chemokine receptor 4 (CXCR4)-encoding cDNA.
DE
XX Human; CXc chemokine receptor 4; CXCR4; CXCR4 inhibitor;
KW small cell lung cancer; SCLC; cellular adhesion; cellular proliferation;
KW metastasis; motility; morphological change; drug screening; monitoring;
KW cytoskeletal; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 89..1147
FT /*tag= a
FT /product= "Human CXCR4"
XX WO2003079020-A2.
XX 25-SEP-2003.
XX 20-MAR-2003; 2003WO-EP002916.
XX 20-MAR-2002; 2002US-0366370P.
XX (DAND) DANA FARBER CANCER INST INC.
XX (NOVS) NOVARTIS AG.
XX Salgia R;
XX WPI; 2003-757037/71.
DR P-PSDB; ADC98646.
XX Use of a CXc chemokine receptor 4 (CXCR4) inhibitor, for modulating
PT cellular adhesion or inhibiting proliferation, movement, or morphological
PT change in a small cell lung cancer (SCLC) cell population, or for
PT treating a subject having SCLC.
XX Disclosure; SEQ ID NO 1; 78pp; English.
PS

XX The invention relates to methods and compositions for inhibiting small
CC cell lung cancer (SCLC) proliferation and metastasis through modulation
CC of the activity or expression of CXc chemokine receptor 4 (CXCR4), and
CC optionally, that of the tyrosine kinase receptor c-Kit. Administration of
CC a CXCR4 inhibitor modulates cellular adhesion and inhibits cellular
CC proliferation, motility, and morphological changes in an SCLC cell.
CC The invention is based on the finding that CXCR4 is
CC ubiquitously expressed, and c-Kit is variably expressed, in SCLC cells.
CC The invention also discloses methods of screening for CXCR4 inhibitors;
CC determining whether a sample of lung cancer cells expresses CXCR4 and is
CC therefore amenable to CXCR4 inhibitor therapy; determining whether a
CC sample of lung cancer cells coexpress CXCR4 and c-Kit is is therefore
CC amenable to treatment with a combination of a CXCR4 inhibitor and a c-Kit
CC inhibitor; and a method of assessing whether CXCR4 inhibitor therapy or
CC CXCR4 inhibitor/c-Kit inhibitor combination therapy should be continued,
CC by determining CXCR4 (and optionally c-Kit) activity in two or more
CC tumour samples taken from an SCLC patient during the course of treatment.
CC The methods of the invention are useful for treating SCLC, for screening
CC for agents for the treatment of SCLC, for determining whether a SCLC from
CC individual patients is suitable for treatment with agents of the
CC invention, and for determining whether such treatment is effective and
CC should be continued. The present sequence represents cDNA encoding human
CC CXCR4, which can be used in methods of the invention.
XX Sequence 1679 BP; 407 A; 400 C; 368 G; 504 T; 0 U; 0 Other;
SQ

Query Match 55.1%; Score 933.4; DB 9; Length 1679;
Best Local Similarity 81.8%; Pred. No. 4.3e-207;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACATCTGTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGATCTCCAAACAG 60
Db 104 ATATACATCTGTGATAACTACTCTGAAGAAGTGGGTCTGGAGACTATGATCTCCAAACAG 163
QY 61 GAACCTCTGCTTCGGGATGAACACCTCCATTTCAATAGATCTTCTGCGCCACCATCTAC 120
Db 164 GAACCTCTGCTTCGGGATGAACACCTCCATTTCAATAGATCTTCTGCGCCACCATCTAC 223
QY 121 TTCATCACTCTTCTGACTGGGATAGTCGCAATGATGTTGATCTCTGTCATGGTTAC 180
Db 224 TCATCATCTCTTCACTGATGTCATTCACGAGGAATGGCTCAGGGACTATGATCTCCATGAG 283
QY 181 CAGAAGAAGCTAAGGAGCATGACGCAAGTACCGGCTGCACCTGTCTAGTGGCTGACCTC 240
Db 284 CAGAAGAAGCTAAGGAGCATGACGCAAGTACCGGCTGCACCTGTCTAGTGGCTGACCTC 343
QY 241 CTCTTTGTCTATCAGCTCCCTTCTGGGAGTTGATGTCATGGCTGACTGGTACTTTGG 300
Db 344 CTCTTTGTCTATCAGCTCCCTTCTGGGAGTTGATGTCATGGCTGACTGGTACTTTGG 403
QY 301 AAATTTTGTGTAAGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC 360
Db 404 AACTTCTATGCAAGGAGTCCATGTCTATACAGTCAACCTCTACAGCAGTGTCTCTC 463
QY 361 ATCTGGCCTTTCATCAGCTGAGCCGTTACCTGCGCATTTGTCACGCGCAACAGTCAA 420
Db 464 ATCTGGCCTTTCATCAGCTGAGCCGTTACCTGCGCATTTGTCACGCGCAACAGTCAA 523
QY 421 AGSCCAAGGAAGTCTGCTGGTGAAGGAGCAGTCTATGTTGGGCTGTGGATCCAGCCCTC 480
Db 524 AGSCCAAGGAAGTCTGCTGGTGAAGGAGCAGTCTATGTTGGGCTGTGGATCCAGCCCTC 583
QY 481 CTCTGACTATACCTGACTTTTCATCTTTTGGCGAGCTGAGCCAGGAGACATCATAGTAGGG 540
Db 584 CTCTGACTATACCTGACTTTTCATCTTTTGGCGAGCTGAGCCAGGAGACATCATAGTAGGG 628
QY 541 GATGACAGGTATCATCTGTGACCGCTTTTACCCGAGTAGCTGTGGATGGTGGTTCAA 600
Db 629 GATGACAGGTATCATCTGTGACCGCTTTTACCCGAGTAGCTGTGGATGGTGGTTCAA 698
QY 601 TTCCAGCATATATGTTGGTGTCTCATCTGCGCGGATGCTCATCTCTCTCTCTTACTGTC 660

Db 589 TTTACGACATCATGGTTGSCCTTATCCCTGCTGGTATTTGTCATCTCTGCTGCTATTGC 748
QY 661 ATCATCATCTCTAAGTGTACACTCAAAGGGCCACAGAGCGCAAGGCCCTCAAGACG 720
Db 749 ATTATCATCTCCAAGTGTACACTCAAAGGGCCACAGAGCGCAAGGCCCTCAAGACC 808
QY 721 ACAGTCATCTCATCTAGCTTTCTTTGCTGCTGCTGCCATATTATGTGGGATCAGC 780
Db 809 ACAGTCATCTCATCTAGCTTTCTTTGCTGCTGCTGCCATATTATGTGGGATCAGC 868
QY 781 ATCGACTCCTTCATCCTTTTGGAGTCATCAAGAGGATGTGACTTCAGAGCATTTGT 840
Db 869 ATCGACTCCTTCATCCTTCCTGGAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG 928
QY 841 CACAAGTGGATCTCCATCAAGAGGCCCTGCTTCTTCCACTGTTGCTGCAACCCCATC 900
Db 929 CACAAGTGGATTTCCATCACAGAGGCCCTGCTTCTTCCACTGTTGCTGCAACCCCATC 988
QY 901 CTCTATGCTCTCTCGGGCCAAAGTTCAAAGCTCTGCCAGCATGCACTCAACTCCATG 960
Db 989 CTCTATGCTTTCTTGGAGCAAATTAACCTCTGCCAGCAGCACTCACCTCTGTG 1048
QY 961 AGCAGAGGCTCCAGCTCAAGATCCCTTCCAAAGGAAAGGGGGTGGACACTTCCGTC 1020
Db 1049 AGCAGAGGCTCCAGCTCAAGATCCCTTCCAAAGGAAAGGGGGTGGACACTTCACTGTT 1108
QY 1021 TCCACGGAGTCAGATCTCCAGTTTCACTCCAGCTAACCCCTTATGCAAGACTTATAT 1080
Db 1109 TCCACTGAGTCTGAGTCTCAAGTTTCACTCCAGCTAA-----CACAGATGTAAA 1160
QY 1081 AATATATATATATATATGATAAGAACTTTTATGTTACACATTTCCAGATATAAGAG 1140
Db 1161 GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTTACACATTTTTCAGATATAAAG 1220
QY 1141 ACTGACCACTTTGTACAGTTTTTTTTTTTTTTTAAATGACTGTTGGGAGTTTATGTTCC 1200
Db 1221 ACTGACCAATATTGTACAGTTTTTTTATGTTGCTGTTGATTTTTGT-----CTTGTGTTTC 1274
QY 1201 TCTAGTTTTTGTAGGTTTGAATTAATTATATAAATAATGTTTTTTTGTGTTTTCATGT 1260
Db 1275 TTTAGTTTTTGTGAAGTTTAAATGACTTATTATA----TAAATTTTTTTTGTTCATAT 1330
QY 1261 GAATGAGCTCTAGGAGGACCTGTGGCAAGTTCTTAGTACTGTTTATCTGTGTGTAG 1320
Db 1331 TGATGTGTCTAGGAGGACCTGTGGCAAGTTCTTAGTACTGTTTATCTGTGTGTAG 1390
QY 1321 GACTGTAGAACTGTAGAGGAAGAACTGAACATTCAGAAATGTGTGTAATTAATAA 1380
Db 1391 GACTGTAGAA-----AAGGGAAGTGAACATTCAGAGCGGTGTAGTGAATCAGCTAAA 1442
QY 1381 GCTAGCGGTGATCTCAGCTGTTGCTGCATA 1411
Db 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

Search completed: May 17, 2004, 06:37:17
Job time : 736.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 05:52:09 ; Search time 135.63 Seconds
(without alignments)
6931.281 Million cell updates/sec

Title: US-09-367-052-7
Perfect score: 1694
Sequence: 1 atatacactctgataacta.....taaaagtcgaataaaactt 1694

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfileseq1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	933.4	55.1	1664	4	US-09-582-224A-5
2	933.4	55.1	1664	4	US-09-023-655-1213
3	933.4	55.1	1679	4	US-09-517-605-14
4	933.4	55.1	1737	1	US-08-202-056-4
5	933.4	55.1	1737	1	US-08-076-093A-3
6	933.4	55.1	1737	1	US-08-701-265-3
7	933.4	55.1	1737	2	US-08-284-586-3
8	933.4	55.1	1737	2	US-08-805-478-3
9	933.4	55.1	1737	2	US-08-802-627A-3
10	933.4	55.1	1737	2	US-08-801-238-3
11	933.4	55.1	1737	2	US-08-801-228-3
12	933.4	55.1	1737	3	US-09-104-296-3
13	933.4	50.5	1737	5	PCT-US94-06380-2
14	855.6	50.5	1225	4	US-09-016-434-1235
15	834	49.2	1317	1	US-08-153-848-45
16	834	49.2	1317	3	US-09-299-843A-45
17	834	49.2	1317	3	US-09-088-337B-45
18	834	49.2	1317	5	PCT-US93-11153-45
19	198.4	11.7	1119	4	US-09-170-496D-65
20	198.4	11.7	1679	1	US-08-202-056-6
21	198.4	11.7	1679	1	US-08-076-093A-5
22	198.4	11.7	1679	1	US-08-701-265-5
23	198.4	11.7	1679	2	US-08-284-586-5
24	198.4	11.7	1679	2	US-08-805-478-5
25	198.4	11.7	1679	2	US-08-802-627A-5
26	198.4	11.7	1679	2	US-08-801-238-5
27	198.4	11.7	1679	2	US-08-801-228-5

28	198.4	11.7	1679	3	US-09-104-296-5
29	198.4	11.7	1679	5	PCT-US94-06380-3
30	198.4	11.7	2818	3	US-08-982-493-7
31	198.4	11.7	2818	3	US-08-628-655-1
32	196.8	11.6	1119	4	US-09-170-496D-199
33	195.8	11.6	2085	3	US-09-299-843A-65
34	195.8	11.6	2085	4	US-09-088-337B-65
35	193.8	11.4	1107	4	US-09-170-496D-19
36	193.8	11.4	1293	4	US-09-016-434-1052
37	193.8	11.4	1670	3	US-08-709-838-1
38	193.8	11.4	1670	3	US-08-829-839-1
39	193.8	11.4	1670	4	US-09-023-655-980
40	193.8	11.4	1670	4	US-09-624-594-1
41	192.2	11.3	1107	4	US-09-170-496D-173
42	186.8	11.0	2751	1	US-08-153-848-23
43	186.8	11.0	2751	3	US-09-299-843A-23
44	186.8	11.0	2751	4	US-09-088-337B-23
45	186.8	11.0	2751	5	PCT-US93-11153-23

ALIGNMENTS

RESULT 1
US-09-582-224A-5
; Sequence 5, Application US/09582224A
; Patent No. 6429308
; GENERAL INFORMATION:
; APPLICANT: IJIMA, Osamu
; APPLICANT: GOTO, Takeshi
; APPLICANT: SHIMADA, Takashi
; TITLE OF INVENTION: HIV Infection Inhibitors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/582,224A
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: PCT/JF99/06534
; PRIOR FILING DATE: 1999/11/24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Microsoft Word
; SEQ ID NO 5
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cDNA of CXCR4
US-09-582-224A-5

Query Match	55.1%;	Score	933.4;	DB	4;	Length	1664;
Best Local Similarity	81.8%;	Pred. No.	1e-210;	Mismatches	216;	Indels	41;
Matches	1154;	Conservative	0;	Gaps	5;		
QY	1	ATATACACTTCTGATACTCTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCCAACAAAG	60				
Db	92	ATATACACTTCTGATACTCTACACCGAGAAATGGCTCAGGGGACTATGACTCCATGAAG	151				
QY	61	GAACCTGCTCCGGGATGAAAAGTCCATTTCAATAGGATCTTCTGCCACCACCTAC	120				
Db	152	GAACCTGCTTTCCTGGAGAAAATGCTAAATTTCAATAAATCTTCTGCCACCACCTAC	211				
QY	121	TTTCATCATCTTCTGACTGGCATAGTCGGCAATGATTTGGTGCCTCGTTCATGGGTAC	180				
Db	212	TCCATCATCTTCTTACTTGGCAATTTGGGCAATGGATTTGGTTCATCTGCTCATGGGTAC	271				
QY	181	CAGAGAAGCTAAGGAGCATACGAGCAAGTACCGGCTGCACCTGTAGTGGGTGACCTC	240				
Db	272	CAGAGAAGCTAAGGAGCATACGAGCAAGTACAGGCTGCACCTGTAGTGGGTGACCTC	331				
QY	241	CTCTTTGTCACACACTCCCTTCTGGGCACTTGCATGCTGGGTGACTGTTCTTGGG	300				
Db	332	CTCTTTGTCACACACTCCCTTCTGGGCACTTGCATGCTGGGTGACTGTTCTTGGG	391				
QY	301	AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGAGGTTCTC	360				

Db 392 AACTTCCTATGCAAGGAGTCCATGTCATCTACACAGTCAACCTCTACAGCAGTGCTC 451
Qy 361 ATCTGCGCTTCATCAGCTGGACCGGTACTCTGCGCATTTGTCACCGCCACCAACAGTCAA 420
Db 452 ATCTGCGCTTCATCAGCTGGACCGGTACTCTGCGCATTTGTCACCGCCACCAACAGTCAA 511
Qy 421 AGGCCAAGGAACCTGCTGGCTGAAGAGCAGTCTATGTGGCGCTCTGGATCCAGCCCTC 480
Db 512 AGGCCAAGGAAGCTGTTGGCTGAAGAGTGTCTATGTGGCGCTCTGGATCCCTGCGCTC 571
Qy 481 CTCCTGACTATACCTGACTTCTATCTTTGCGAGCTCAGCCAGGCGGACATCAGTCAGGG 540
Db 572 CTGCTGACTATTCGAGTCTCATCTTTGCCAAG-----TCAGTAGGCA 616
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTTTCAA 600
Db 617 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTGTGTCCAG 676
Qy 601 TTCAGCATATAATGGTGGGTCTCATCTGCGGATGCTCATCTCTCTCTCTCTCTCTCT 660
Db 677 TTTGAGCATATGTTGGCTTATCTGCTGTGTATTTGATCTCTCTCTCTCTCTCTCTCT 736
Qy 661 ATCATCATCTTAAGCTGTCACTCCAGGGGCCACAGAGCGCAAGGCCCTTCAAGAGC 720
Db 737 ATTATCATCTCAAGCTGTCACTCCAGGGGCCACAGAGCGCAAGGCCCTTCAAGAGC 796
Qy 721 ACAGTCATCTCATCT 780
Db 797 ACAGTCATCTCATCT 856
Qy 781 ATCGACTCTCTCATCT 840
Db 857 ATCGACTCTCTCATCT 916
Qy 841 CACAAGTGATCTCATCATCAGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 917 CACAAGTGATTTCCATCAGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 976
Qy 901 CTCCTATGCT 960
Db 977 CTCCTATGCT 1036
Qy 961 AGCAGAGCTCAGCTCAGATCT 1020
Db 1037 AGCAGAGGCTCAGCTCAGATCT 1096
Qy 1021 TCCAGGAGTCAGATCT 1080
Db 1097 TCCAGTGTCTGAGTCTTCAAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1148
Qy 1081 AAT 1140
Db 1149 GACTTTTTTTTATACGATAAATAAATTTTTTTTAAAGTTTACATATTTTTCAGATATAAAG 1208
Qy 1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTATGTTCC 1200
Db 1209 ACTGACCAATATGTACAGTTTTTTTATGCTTGTGGATTTTTTGT-----CTTGTGTTTC 1262
Qy 1201 TCTAGTTTTTGTGAGTTTGAATTTATATATATATATATATATATATATATATATATAT 1260
Db 1263 TTTAGTTTTTGTGAAGTTTAAATGACTTATTTATA-----TAAATTTTTTTTGTTCATAT 1318
Qy 1261 GAATGAGGCTTAGCAGGACTGTGGCAAGTTCTTAGTAGCTGTGTATCTGTGTGTAG 1320
Db 1319 TGATGTGTCTAGGAGGAGCTGTGGCAAGTTCTTAGTTGCTGTGTGTGTGTGTGTGTGT 1378
Qy 1321 GACTGTAGAACTGTAGAGGAAGAACTGAAATCCAGAAATGTGTGGTAAATGAATAAA 1380
Db 1379 GACTGTAGNA-----AGGGAAGTGAATTTCCAGAGGCTGTAGTTAATCAGGTAAA 1430
Qy 1381 GCTAGCGGTGATCTCAGCTGT 1411
Db 1431 GCTAGAAATGATCCCGCAGCTGT 1461

RESULT 2

US-09-023-655-1213
; Sequence 1213, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g219868
; US-09-023-655-1213

Query Match 55.1%; Score 933.4; DB 4; Length 1664;
Best Local Similarity 81.8%; Pred. NO. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACATCTTGATAACTCTCTGAAGAAGTGGGCTGGAGACTATGACTCCCAACAG 60
Db 92 ATATACATCTTGATAACTCTACACCGAGAAATGGGCTCAGGAGACTATGACTCCCAAG 151
Qy 61 GAACCTCTGCTCCGGGATGAAGAACTCCATTTCAATAGGATCTTCTCCCAACCATCTAC 120
Db 152 GAACCTCTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCTCCCAACCATCTAC 211
Qy 121 TTTCATCATCTTTTACTGCTGGCATAGTCGGCATGATTTGGTGTATCTGTCATGGTTAC 180
Db 212 TCCATCATCTTTTAACTGGCATTTGTTGGCAATGATTTGGTGTATCTGTCATGGTTAC 271
Qy 181 CAGAAGAACTAAGGAGCATGACGACAAAGTACCGGCTGCACCTCTCAGTGGCTGACCTC 240
Db 272 CAGAAGAACTAAGGAGCATGACGACAAAGTACAGGCTGCACCTCTCAGTGGCTGACCTC 331
Qy 241 CTCCTTGTCTACACCTCCCTCTTGGGCAAGTGTATGCAATGGCTGACTGTGTTGGG 300

Db 332 CTCTTTGTGATCAGGCTTCCCTTCTGGGAGTTGATGCGGTGGCAAACTGGTACTTTGGG 391
Qy 301 AAATTTTGTGTAAGGCTGTCATATCACTGTCACCTCTACAGAGCGTTCTC 360
Db 392 AACTTCTATGCAAGGAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGCCTC 451
Qy 361 ATCTGCGCTTCTATCAGCTGGAGCGGTACTCGCCANTGTCCACGCCCAACAGTCAA 420
Db 452 ATCTGCGCTTCTATCAGCTGGAGCGGTACTCGCCANTGTCCACGCCCAACAGTCAA 511
Qy 421 AGGCCAAGGAACTGCTGGCTGAAAGGAGTCTATGTGGGCGTCTGGATGCCAGCCCTC 480
Db 512 AGGCCAAGGAGTGTGGCTGAAAGGAGTCTATGTGGGCGTCTGGATGCCAGCCCTC 571
Qy 481 CTCCTGACTATCTGACTTCACTTTGCGAGTGCAGCCAGGGGAGCATCAGTCAAGGG 540
Db 572 CTGCTGACTATCCCGACTTCTATCTTTGCCAAGC-----TCAGTAGGCA 616
Qy 541 GATCAGAGTACATCTGTGACCGCTTTACCCGATAGCCTGTGGATGGTGTTCAA 600
Db 617 GATCAGAGATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGTGTGTCTCCAG 676
Qy 601 TTCAGCATATATGTGGGTCTCATCTGCGCGGATCTCATCTCTCTCTCTCTCTCTCT 660
Db 677 TTTTCAGCATATCTGTGGCTTATCTCTGCTGTATTTGTCTATCTCTCTCTCTCTCTCT 736
Qy 661 ATCATCTCTCTAGCTGTCACTCTCAAGGCGCCACAGAGCGCAAGGCGCCCTCAAGAG 720
Db 737 ATTATCATCTCAGAGTGTCACTCTCAAGGCGCCACAGAGCGCAAGGCGCCCTCAAGAG 756
Qy 721 ACAGTCACTCTCATCTGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 797 ACAGTCACTCTCATCTGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
Qy 781 ATCGACTCTCTCATCTTCTGAGTCTATCAAGCAAGATGTGACTTTCGAGAGCATGTG 840
Db 857 ATCGACTCTCTCATCTTCTGAGTCTATCAAGCAAGATGTGACTTTCGAGAGCATGTG 916
Qy 841 CACAAGTGGATCTCATCAGAGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 917 CACAAGTGGATTTCCATCAGGAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Qy 901 CTCTATGCTTCTCGGGGCGAGTTCAGAGCTCTGCCAGCATGCACTCAACTCCCATG 960
Db 977 CTCTATGCTTCTCGGGGCGAGTTCAGAGCTCTGCCAGCATGCACTCAACTCCCATG 1036
Qy 961 AGCAGAGCTCTCAGCTCAAGATCTTTCAGAGGAAAGCGGGGTGGACACTCTTCCGTC 1020
Db 1037 AGCAGAGGCTCCAGCTCAAGATCTTTCAGAGGAAAGCGGGGTGGACACTCTTCCGTC 1096
Qy 1021 TCCAGGAGTCAAGATCTCAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1097 TCCAGTGAAGTCTTCAAGTCTTCAAGTCTTCACTCCAGCTAA-----CACAGATGTAAA 1148
Qy 1081 AAT 1140
Db 1149 GACTTTTTTTTATAGATATAACTTTTTTTTAAAGTAAAGTAAAGTAAAGTAAAG 1208
Qy 1141 ACTGACAGTCTGTGACAGTTTTTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTAGTTC 1200
Db 1209 ACTGACCAATATGTGACAGTTTTTTTATGCTGTGTGGATTTTGT-----CTTGTGTTT 1262
Qy 1201 TCTAGTTTTTGTGAGTTGACTTAATATATATATATATATATATATATATATATATAT 1260
Db 1263 TTTAGTTTTTGTGAGTTTAAATGACTTATTTATA-----TAAATTTTTTTTGTTCATAT 1318
Qy 1261 GAATGAGGCTCTAGGAGGAGCTGTGGCAAGTCTTCTAGTACTGTCTTATCTGTGTAG 1320
Db 1319 TGAATGTGTCTAGGAGGAGCTGTGGCAAGTCTTCTAGTACTGTCTTATCTGTGTAG 1378
Qy 1321 GACTGTAGAACTGTAGAGAGAAACTGAATTCAGAAATGTGTGTGTGTGTGTGTGTGTGT 1380
Db 1379 GACTGTAGAA-----AAGGAACTGAACATTCAGAGCGGTGTAGTTATTCAGCTAAA 1430

Qy 1381 GCTAGCCGTGATCTCTCAGCTGTCTGCTGCATA 1411
Db 1431 CTTAGAAATGATCCCGAGCTGTTTATGCATA 1461
RESULT 3
US-09-517-605-14
; Sequence 14, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-517-605-14
Query Match 55.1%; Score 933.4; DB 4; Length 1679;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
Qy 1 ATATACACTTCTGATAACTACTCTGAAAGAGTGGGCTGTGGAGACTATGACTCCCAACAG 60
Db 104 ATATACACTTCTGATAACTACTACCCAGGAAATGGGCTGTGGAGACTATGACTCCCAACAG 163
Qy 61 GAACCTCTGCTCCGGATGAAACGTCATTTCAATAGAGATCTTCTGCCCCACCATCTAC 120
Db 164 GAACCTCTGCTCCGGATGAAACGTCATTTCAATAGAGATCTTCTGCCCCACCATCTAC 223
Qy 121 TTATCATCTTCTTGAATGAGTGTGGCAATGATTTGGTGTATGCTGTATGCTGTATG 180
Db 224 TCCATCATCTTCTTGAATGAGTGTGGCAATGATTTGGTGTATGCTGTATGCTGTATG 283
Qy 181 CAGAGAACTAAGAGAGTACGAGCAAGTACCGGCTGACCTGTGCTGCTGCTGCTGCTGCT 240
Db 284 CAGAGAACTAAGAGAGTACGAGCAAGTACCGGCTGACCTGTGCTGCTGCTGCTGCTGCT 343
Qy 241 CTCTTTGTCTATCAGCTCTCCCTTCTGGGCAATGATTTGGTGTATGCTGTATGCTGTATG 300
Db 344 CTCTTTGTCTATCAGCTCTCCCTTCTGGGCAATGATTTGGTGTATGCTGTATGCTGTATG 403
Qy 301 AAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCAACCTCTACAGAGCGTTCTC 360
Db 404 AAATTTTGTGTAAGGCTGTCATATCATCTACACTGTCAACCTCTACAGAGCGTTCTC 463
Qy 361 ATCTGCGCTTCTATCAGAGTGGAGCGGTACTCGCCANTGTCCACGCCCAACAGTCAA 420
Db 464 ATCTGCGCTTCTATCAGTCTGGAGCGGTACTCGCCANTGTCCACGCCCAACAGTCAA 523
Qy 421 AGGCCAAGGAACTGCTGGCTGAAAGGAGTCTATGTGGGCGTCTGGATGCCAGCCCTC 480
Db 524 AGGCCAAGGAACTGCTGGCTGAAAGGAGTCTATGTGTGGGCGTCTGGATGCCAGCCCTC 583
Qy 481 CTCTGACTATATCTGACTTCTATCTTTGCGAGCTCAGCCAGGGGAGCATCAGTCAAGGG 540
Db 584 CTCTGACTATATCTGAGCTTCTATCTTTGCCAAGC-----TCAGTAGGCA 628
Qy 541 GATGACAGTACATCTGTGACCGCTTTTACCCCGATAGCCTGTGATGCTGTGTGTCTCA 600
Db 629 GATGACAGATATATCTGTGACCGCTTCTTACCCCAATGACTTTGTGGGTGTGTGTGTCT 688
Qy 601 TTCCAGCATATATGTTGGGTCTCATCTCTGCGGCACTGCTCATCTCTCTCTCTCTCTCT 660

601 TTCCAGCATATATGTTGGTCTCTCATCTGCCGGCATCGTCTCTCTCTGTTACTGC 660
691 TTTACACATCATGTTGGCTTATCCCTGCTGTATGTTCTCTCTCTGTTACTGC 750
661 ATCATCATCTTAAGCTGTACATCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 720
751 ATTATCATCTCAAGCTGTACATCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 810
721 ACAGTCACTCTCATCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
811 ACAGTCACTCTCATCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
781 ATGAGCTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
871 ATGAGCTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930
841 CACAAGTGTATCTCAATCATCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 900
931 CACAAGTGTATCTCAATCATCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 990
901 CTCTATGCT 960
991 CTCTATGCT 1050
961 AGCAGAGGCTCTCAAGCTCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 1020
1051 AGCAGAGGCTCTCAAGCTCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 1110
1021 TCCAGGAGTCAAGCTCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 1080
1111 TCCAGGAGTCAAGCTCTCAAGGCGCCAGGAGCGCAAGCCCTCAAGAG 1162
1081 AAT 1140
1163 GACTTTTTTATACGATAATAAATATATATATATATATATATATATATATAT 1222
1141 ACTGACGAGTCTGTACAGTTTTTTTTTTTTTTTTTTTTTTTTTAAATGAGT 1200
1223 ACTGACGAGTCTGTACAGTTTTTTTTTTTTTTTTTTTTTTTTTAAATGAGT 1276
1201 TCTAGTTTTTGTGAGGTTTGACTTAAATATATATATATATATATATATATAT 1260
1277 TTTAGTTTTTGTGAGTAAATGACTTATATATA-----TAAATTTTTTTT 1332
1261 GAATGAGGCTCTAGGAGGAGCTGTGGCCAAAGTTCTTAGTAGTGTATCTGTGTAG 1320
1333 TGATGTGTCTAGGAGGAGCTGTGGCCAAAGTTCTTAGTAGTGTATCTGTGTAG 1392
1321 GACTGTAGAACTGTAGAGGAGAACTGAACTTCCAGAGTGTGTGTAATTAATA 1380
1393 GACTGTAGAA-----AAGGGAAGTGAACATTCAGAGCGGTGTAGTGAATCA 1444
1381 GCTAGCGCTGATCTCAGCTGTGTGTCGATA 1411
1445 GCTAGAAATGATCCAGCTGTGTATGATA 1475

RESULT 5

US-08-076-093A-3
; Sequence 3, Application US/08076093A
; Patent No. 5543503
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4A Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A
FILING DATE: 11-Jun-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-076-093A-3

Query Match 55.1%; Score 933.4; DB 1; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTTCTGATTAATACTCTGAAAGAGTGGGTCTGGAGACTATGATCCCAAG 60
DB 106 ATATACACTTCTGATTAATACTCTGAAAGAGTGGGTCTGGAGACTATGATCCCAAG 165
QY 61 GAACCTCTCTTCCGGGATGAAACGTCATTTCAATAGGATCTTCTGCCACCATCTAC 120
DB 166 GAACCTCTCTTCCGGGATGAAACGTCATTTCAATAGGATCTTCTGCCACCATCTAC 225
QY 121 TTCTATCATCTTCTGACTGTGGCATAGTGGCAATGGATTTGGTATCTCTGGTATCGGTTAC 180
DB 226 TCCATCATCTTCTTAACTGGCATTTGGGCAATGGATTTGGTATCTCTGGTATCGGTTAC 285
QY 181 CAGAGAGCTTAAGAGCATGACGAGCAAGTACCGGCTGCACCTGTGATGGCTGACCTC 240
DB 286 CAGAGAGCTTAAGAGCATGACGAGCAAGTACCGGCTGCACCTGTGATGGCTGACCTC 345
QY 241 CTCTTTGTCATCACATCTCTCTGGGCAATGGATTTGGTATGGCTGATCTTTGGG 300
DB 346 CTCTTTGTCATCACATCTCTCTGGGCAATGGATTTGGTATGGCTGATCTTTGGG 405
QY 301 AAATTTTGTGTAGGCTGTCCATATCATCTACACTGTCACTCTACAGCAGGTTCTC 360
DB 406 AACTTCTTATGAAGGAGTCCATGTATCTACACAGTCACTCTACAGCAGTGTCTC 465
QY 361 ATCTGGCTTTTCATCAGCTCTGAGCGGTACCTGCCATTTGCCACGCCACCAAGTCAA 420
DB 466 ATCTGGCTTTTCATCAGCTCTGAGCGGTACCTGCCATTTGCCACGCCACCAAGTCAA 525
QY 421 AGGCCAAGAACTGTGGCTGAAAGAGCAGTCTATGTGGGCGTCTGATCCAGCCCTC 480
DB 526 AGGCCAAGAACTGTGGCTGAAAGAGCAGTCTATGTGGGCGTCTGATCCCTGCCCC 585
QY 481 CTCTGTACTTACTGACTTCTCTTGGCCAGCTCAGCCAGGGGAGCATCAGTCAGGGG 540
DB 586 CTGCTGACTATTTCCGACTTCTATCTTTGCCAAG-----TCAGTGAGGCA 630

541 GATGACAGGTACATCTGTGACCCGCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
Db GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTGGTGTTCAG 690
Qy TTCAGCATATATGGTGGTCTCATCTGCGGGCATGTCATCTCTCTCTCTCTCTCTCTCT 660
Db TTCAGCATATATGGTGGTCTCATCTGCGGGCATGTCATCTCTCTCTCTCTCTCTCTCT 750
Qy ATCATCATCTCTAAGTGTGACACTCTCAAGGGGCCACAGAGGCGCAAGGCGCTTCAAGAG 720
Db ATTATCATCTCCAGCTGTACACTCTCAAGGGGCCACAGAGGCGCAAGGCGCTTCAAGAG 810
Qy ACAGTCACTCTCATCT 780
Db ACAGTCACT 870
Qy ATGACCT 840
Db ATGACCT 930
Qy CACAGTGGATCTCCATCAAGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db CACAGTGGATCTCCATCAAGAGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 990
Qy CTCATGCT 960
Db CTCATGCT 1050
Qy AGCAGAGGCTCCAGCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db AGCAGAGGCTCCAGCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1110
Qy TCCAGGAGTCAAGTCT 1080
Db TCCAGGAGTCAAGTCT 1162
Qy AAT 1140
Db GACTTTTTTTTACGATAAATAAATATATATATATATATATATATATATATATATAT 1222
Qy ACTGACAGTCTGTACAGTCT 1200
Db ACTGACAGTCTGTACAGTCT 1276
Qy TCTAGTTTTTGTGAGGTTTCACTTAATTTATATATATATATATATATATATATATAT 1260
Db TTTAGTTTTTGTGAGGTTTCACTTAATTTATATATATATATATATATATATATATAT 1332
Qy GAATGAGCTGTAGGAGGACCTGTGCGCAAGTTCTTAGTGTGTGTGTGTGTGTGTGTGT 1320
Db TGATGTGTGTGTAGGAGGACCTGTGCGCAAGTTCTTAGTGTGTGTGTGTGTGTGTGT 1392
Qy GACTGTAGAACTGTAGAGGAGAACTGTAACTTCCAGAAATGTGTGTGTGTGTGTGTGT 1380
Db GACTGTAGAA-----AAGGAACTGTAACTTCCAGAAATGTGTGTGTGTGTGTGTGT 1444
Qy GCTAGCGGTGATCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1411
Db GCTAGAAATGATCCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1475

RESULT 6

US-08-701-265-3
; Sequence 3, Application US/08701265
; Patent No. 5776457
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: Antibodies to Human PF4 Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WisPatIn (Genentech)
CURRENT APPLICATION DATA: US/08/701,265
APPLICATION NUMBER: 08/076093
FILING DATE: 11-Jun-1993
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-701-265-3

Query Match 55.1%; Score 933.4; DB 1; Length 1737;

Best Local Similarity 81.8%; Pred. No. 18-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACACTTCTGATAACTACTCTGAAGAAAGTGGGTCTGGAGACTATGACTCCAAAG 60
Db 106 ATATACACTTCTGATAACTACTACCGAGAAATGGCTCAGGGACTATGACTCCATGA 165
Qy 61 GAACCTGTCTCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGGACCATCTAC 120
Db 166 GAACCTGTCTCCGGGATGAAAGCTCCATTTCAATAGGATCTTCTGCGGACCATCTAC 225
Qy 121 TCCATCATCTTCTTGAAGTGTGCGGATGCGCAATGATGCGGATGCGGATGCGGATG 180
Db 226 TCCATCATCTTCTTGAAGTGTGCGGATGCGCAATGATGCGGATGCGGATGCGGATG 285
Qy 181 CAGAAGAGCTTAAGAGCATGACGCAAGTACCGGCTGCACTGTCACTGTCACTGTGAGCTC 240
Db 286 CAGAAGAGCTTAAGAGCATGACGCAAGTACCGGCTGCACTGTCACTGTCACTGTGAGCTC 345
Qy 241 CTCTTTGTCTACATCCCTCTCTGCGGATGATGCGGATGCGGATGCGGATGCGGATG 300
Db 346 CTCTTTGTCTACATCCCTCTCTGCGGATGATGCGGATGCGGATGCGGATGCGGATG 405
Qy 301 AAATTTTGTGTAAGGTGTCCATATCATCTACACTGTCAACCTCTACAGCAGGTTCTC 360
Db 406 AACTTCTATGCAAGGCGAGTCCATGTCTATACAGCTCAACCTCTACAGCAGGTTCTC 465
Qy 361 ATCTGGCTTCTACAGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 420
Db 466 ATCTGGCTTCTACAGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCG 525
Qy 421 AGGCAAGAGAACTGTGGGTGAAAGGAGCTATGTGGGCGCTGTGGATCCCGAGCTC 480

526 AGCCAAAGGAGCTGTTGGCTGAAGAGGTGCTATGTTGGGCTGTGGATCCCTGCCCTC 585
Db
481 CTCTGACTATACCTGACTCATCTTTTGGAGCTCAGCCAGGGGACATCATGTCAGGG 540
Qy
586 CTGCTGACTATTCCTGACTCATCTTTTGGCAAG-----TCAGTGAGGCA 630
Db
541 GATGACAGGTACATCTGTGACCCGCTTTACCCGATAGCTGTGGATGGTGTGTTCAA 600
Qy
631 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTGTGTTCCAG 690
Db
601 TTCAGCATATATAGTGGGTCTCATCTGCGCGGCACTGTCATCTCTCTGTTACTGC 660
Qy
691 TTTTCAGCATATATAGTGGGTCTCATCTGCGCTGTATGTCATCTCTCTGTTACTGC 750
Db
661 ATCATCATCTTAAGCTGTACACTCAAGGGCCACAGAGGCGCAAGGCGCTCAAGAG 720
Qy
751 ATTATCATCTCCAGCTGTGACCTCTCGAAATCATCAAGAGGCGCAAGGCGCTCAAGACC 810
Db
721 ACAGTCATCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy
811 ACAGTCATCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
Db
781 ATGACTCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy
871 ATGACTCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
Db
841 CACAAGTGGATCTCCATCAAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy
931 CACAAGTGGATCTCCATCAAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
Db
901 CTCTATGCTCTCTCGGGGCAAGTCTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy
991 CTCTATGCTCTCTCGGGGCAAGTCTCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
Db
961 AGCAGAGGCTCCAGGCTCAAGATCTCTTCCAAAGGAAAGCGGGGTGAGCACTCTTCCGTC 1020
Qy
1051 AGCAGAGGCTCCAGGCTCAAGATCTCTTCCAAAGGAAAGCGGGGTGAGCACTCTTCCGTC 1110
Db
1021 TCCAGGAGTCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy
1111 TCCAGTCAAGTCTGAGTCTCAAGTCTTCTCAAGTCTTCTCAAGTCTTCTCAAGTCTTCTCAAG 1162
Db
1081 AAT 1140
Qy
1163 GACTTTTTTTTACGTAATAAATCACTTTTTTTTAAAGTACACATTTTTCAGATATATAAG 1222
Db
1141 ACTGACCAAGTCTGACAGTTTTTTTTTTTTTTTTTTTAAATGACTCTGCTGCTGCTGCTGCTGCT 1200
Qy
1223 ACTGACCAATATGACAGTTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276
Db
1201 TCTAGTTTTTTGAGGTTGACTTAATTTATATATAATTTTGTGTTTTTTGTTTTTTCATGT 1260
Qy
1277 TTTAGTTTTTTGAGGTTGACTTAATTTATATATAATTTTGTGTTTTTTGTTTTTTCATGT 1332
Db
1261 GAATGAGGCTCTAGGACGAGCTGTGCGCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
Qy
1333 TGATGTGCTAGGACGAGCTGTGCGCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1392
Db
1321 GACTGTAGAACTGTAGAGGAGAACTGAACATTCAGATGTGTGGTAAATTTGAATATA 1380
Qy
1393 GACTGTAGAA-----AAGGGAACCTGAACATTCAGATGTGTGGTAAATTTGAATATA 1444
Db
1381 GCTAGCGGTGATCTCAGCTGTTTCTGTCATA 1411
Qy
1445 GCTAGAAATGATCCAGCTGTTTATGATA 1475
Db

RESULT 7

US-08-284-586-3

; Sequence 3, Application US/08284586

; Patent No. 5840856

; GENERAL INFORMATION:

; APPLICANT: Chuncharapai, Anan

APPLICANT: Lee, James
APPLICANT: Hebert, Caroline
APPLICANT: Jin Kim, K.
TITLE OF INVENTION: Antibodies to Human PF4A Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,586
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/076,093A
FILING DATE: 11-Jun-1993
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/677211
FILING DATE: 29-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: 706P2
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
LENGTH: 1737 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-284-586-3

Query Match 55.1%; Score 933.4; DB 2; Length 1737;

Best Local Similarity 81.8%; Pred. No. 1e-210;

Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACACTTCTGATACTACTCTGAAGAGTGGGTCTGGAGACTATGACTCCAAACAAG 60
Db 106 ATATACACTTTCAGATAACTACACCGAGGAATGGCTCAGGGGACTATGACTCCATGAAG 165
Qy 61 GAACCTGCTTCCGGGATGAAGACGTCCATTTCAATAGGATCTTCTGCCCAACCATCTAC 120
Db 166 GAACCTGCTTCCGGGATGAAGAAATGCTAAATTCATATAAATCTTCTGCCCAACCATCTAC 225
Qy 121 TTTCATCATCTTCTTCACTGGCATAGTCGGCAATGATGGTGGTATGCTGCTGCTGCTGCTGCT 180
Db 226 TCCATCATCTTCTTCACTGGCATAGTCGGCAATGATGGTGGTATGCTGCTGCTGCTGCTGCT 285
Qy 181 CAGAGAGAGCTAAGGAGCATGACGACAAAGTACCGGCTGACCTGTGTAGTGGGTGACCTC 240
Db 286 CAGAGAGAACTGAGAGCATGACGACAAAGTACCGGCTGACCTGTGTAGTGGGTGACCTC 345
Qy 241 CTCCTTGTGATCAGACTCCCTCTTGGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 346 CTCCTTGTGATCAGACTCCCTCTTGGGAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 405
Qy 301 AAATTTTGTGTAGGCTGTCCATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 406 AACTTCTCTATGCAAGGAGTCCATGTCTATACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 465

QY 361 ATCTGCGCTTCATCAGCTGGACCGGTACCTCGCCATTTGTCCACGCCACCAACAGTCAA 420
Db 466 ATCTGCGCTTCATCAGCTGGACCGGTACCTCGCCATTTGTCCACGCCACCAACAGTCAA 525
QY 421 AGCCCAAGGAACTGCTGGCTGAAAAGGAGTCTATGTGGGGCTCTGGATCCAGCCCTC 480
Db 526 AGCCCAAGGAACTGCTGGCTGAAAAGGAGTCTATGTGGGGCTCTGGATCCAGCCCTC 585
QY 481 CTCTGAGTATACCTGACTTCATCTTTGGCGAGTCCAGCCAGGGGAGATCATGTCAGGG 540
Db 586 CTCTGAGTATACCTGACTTCATCTTTGGCGAGTCCAGCCAGGGGAGATCATGTCAGGG 630
QY 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGTGTTCAA 600
Db 631 GATGACAGGTATATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGTGTTCAA 690
QY 601 TTCAGCATATATGTTGGGTCTCATCTCTGCGCGGATCGTATCTCTCTGTTACTGC 660
Db 691 TTTTCAGCATATGTTGGGTCTCATCTCTGCGCGGATCGTATCTCTCTGTTACTGC 750
QY 661 ATCATCATCTTAAGCTGTACACTCAAGGGCCACAGAGGCGCAAGGCCCTCAAGAG 720
Db 751 ATTATCATCTCAAGCTGTACACTCAAGGGCCACAGAGGCGCAAGGCCCTCAAGAG 810
QY 721 ACAGTCATCTCATCTGAGCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 811 ACAGTCATCTCATCTGAGCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
QY 781 ATGAGCTCTCTCTCTTTTGGGAGTCATCAAGCAAGGATGACTTCGAGAGCATGAG 840
Db 871 ATGAGCTCTCTCTCTTTTGGGAGTCATCAAGCAAGGATGACTTCGAGAGCATGAG 930
QY 841 CACAAGTGGATCTCCATCAGAGGGCCCTCGCTCTCTTCCACTGTTGCTGAAACCCATC 900
Db 931 CACAAGTGGATCTCCATCAGAGGGCCCTCGCTCTCTTCCACTGTTGCTGAAACCCATC 990
QY 901 CTCTATGCTCTCTCGGGGCAAGTTCAAAGCTCTGCGCCAGCATGCTCACTCACTCATG 960
Db 991 CTCTATGCTCTCTCGGGGCAAGTTCAAAGCTCTGCGCCAGCATGCTCACTCACTCATG 1050
QY 961 AGCAGAGGCTCCAGCTTCAAGTCTTTTCCAAAGGAAAGCGGGGTGAGCACTCTCCGTC 1020
Db 1051 AGCAGAGGCTCCAGCTTCAAGTCTTTTCCAAAGGAAAGCGGGGTGAGCACTCTCCGTC 1110
QY 1021 TCACAGGAGTCAAGTCTCCAGTCTTCTCAGCTTAACCCCTTATGCAAGACTTATAT 1080
Db 1111 TCACAGTGAAGTCTGAGTCTTCAAGTCTTCTCAGCTTCTCAGCTTCTCAGCTTCTG 1162
QY 1081 AATATATATATATATGATAAGAACTTTTATGTTACACATTTTCCAGATATAAG 1140
Db 1163 GACTTTTTTTTACGATAAATACTTTTTTTTAAAGTTACACATTTTCCAGATATAAG 1222
QY 1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGTGGGAGTTTATGTTCC 1200
Db 1223 ACTGACCAATATTGTACAGTTTTTTTATGCTGTTGTTGATTTTTGT-----CTTGTGTTTC 1276
QY 1201 TCTAGTTTTTGTGAGTTTCACTTAATTTATATAATATGTTTGTGTTTGTGTTTGTGTT 1260
Db 1277 TTTAGTTTTTGTGAGTTTAAATGACTTATTTATA-----TAAATTTTTTTTGTGTTTCAATAT 1332
QY 1261 GAATGACGCTGTAGGAGGACCTGTGGCCAAAGTCTTATGAGTGTGTTTATCTGTGTGAG 1320
Db 1333 TGATGTGTGTAGGAGGACCTGTGGCCAAAGTCTTATGAGTGTGTTTATCTGTGTGAG 1392
QY 1321 GACTGTAGAACTGTAGAGGAGAACTGAAACATTCAGAAATGTGTGTTGTTAAATGATAAA 1380
Db 1393 GACTGTAGAA-----AAGGGAACCTGAAACATTCAGAGGCGGTGTAGTGAATCAGTAAA 1444
QY 1381 GCTAGCGGTGATCTCAGCTGTTGCTGCATA 1411
Db 1445 GCTAGAAATGATCCCAAGCTGTTTATGCTATA 1475

RESULT 8
US-08-805-478-3
; Sequence 3, Application US/08805478
; Patent No. 5874543
; GENERAL INFORMATION:
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: Jin Kim, K.
; TITLE OF INVENTION: ANTIBODIES TO PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,478
; FILING DATE: 25-Feb-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-805-478-3

Query Match 55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCCAACAAG 60
Db 106 ATATACACTCTGATAACTACTCAGCAGGAATGGGCTCAGGGGACTATGACTCCATGAAG 165
QY 61 GAACCCCTGTTCCGGGATGAAAACGTCCATTTCAATAGGATCTTCTCGCCACCATCTAC 120
Db 166 GAACCCCTGTTCCGGTGAAGAAAATGCTAAATTTCAATAAAATCTTCTCGCCACCATCTAC 225
QY 121 TTCATCATCTCTTCTGACTGGCATAGTCGCAATGGATTGGTGTGATCTCTGCTCATGGTTAC 180
Db 226 TCCATCATCTCTTCTTAACTGGCAATTTGGGCAATTTGGTTCATCTCTGCTCATGGTTAC 285
QY 181 CAGAAGAAGCTAAGAGCATGACGCAAGTACCGGCTGCACTGTGAGTGGTGTGACCTC 240
Db 286 CAGAAGAAGCTAAGAGCATGACGCAAGTACAGGCTGCACTGTGAGTGGTGTGACCTC 345
QY 241 CTCTTTGTATCATCACCTCCCTTCTGGGCAAGTTGATGCCATGGCTGACTGTTGGG 300

346 CTCTTTGTATCAGCTTCCCTTCTGGGAGTGTGATCGGTGGCAACTGGTACTTTGGG 405
301 AAATTTTGTGTAGGTGTGTCATATCATCTACATCTGTCAACCTCTACACAGAGTTCTC 360
406 AACTTCTCTATGCAAGGAGTCCATCTCATCTACACAGTCAACCTCTACACAGAGTCTC 465
361 ATCTGGCTTCTATCAGCTGAGCGGTACTCTGCCATTTGTCACGCGCCACCAACAGTCAA 420
466 ATCTGGCTTCTATCAGTCTGACCGCTACTCTGCCATCTGTCACGCGCCACCAACAGTCAA 525
421 AGGCAAGGAACTGTCTGGCTGAAAGAGCGAGTCTATGTGGGCGTCTGGATCCACGCGCTC 480
526 AGGCAAGGAACTGTCTGGCTGAAAGAGTGTCTATGTGGGCGTCTGGATCCCTGCGCTC 585
481 CTCTGACTATACCTGACTTCTATCTTTCGCGAGTCTGACGCGGAGGAGCATCAGTCAAGGG 540
586 CTGCTGACTATTCGAGCTTCTATCTTTCGCAAG-----TCAGTGAGGCA 630
541 GATGACAGGTATCTGTGACCGCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
631 GATGACAGATATCTGTGACCGCTTTACCCCAATGACTGTGGTGGTGTGTTCAG 690
601 TTCAGCATATATGTTGGGTCTCATCTGCGCGCATCTGTCATCTCTCTGTACTGC 660
691 TTTGAGCATATCTGTGAGCGCTTATCTGCTGTGTATTTGTCCTCTGTCTGTATTC 750
661 ATCATCATCTTAAGCTGTACATCTCAAGGCGCCACAGAGCGCAAGGCGCTCAAGAGC 720
751 ATTATCATCTCAAGCTGTACATCTCAAGGCGCCACAGAGCGCAAGGCGCTCAAGAGC 810
721 ACAGTCATCTCATCTGAGTCTTCTTTCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 780
811 ACAGTCATCTCATCTGAGTCTTCTTTCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 870
781 ATCGACTCTTCTCATCTTTCGAGTCTATCAAGAGGATGTGACTTCGAGAGCATTTGTG 840
871 ATCGACTCTTCTCATCTTTCGAGTCTATCAAGAGGATGTGACTTCGAGAGCATTTGTG 930
841 CACAAGTGGATCTCATCAAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
931 CACAAGTGGATCTCATCAAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
901 CTCTATGCTTCTCGGGGCGAAGTTCAAAAGCTCTGCGCCAGCATGCACTCAACTCCATG 960
991 CTCTATGCTTCTCGGGGCGAAGTTCAAAAGCTCTGCGCCAGCATGCACTCAACTCCATG 1050
961 AGCAGAGGCTCCAGCTCAAGATCTTTCCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1051 AGCAGAGGCTCCAGCTCAAGATCTTTCCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1110
1021 TCCAGGAGTCAAGATCTTTCCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1111 TCCAGTGTGAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 1162
1081 AAT 1140
1163 GACTTTTTTTATATAGATATAAATACTTTTTTTTAAAGTTACATTTTTTACATATAAAG 1222
1141 ACTGACAGTCTGTGACAGTTTTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTAATGTC 1200
1223 ACTGACCAATATGACAGTTTTTTTTATGCTGTGTTGATTTTGT-----CTTGTTTC 1276
1201 TCTAGTTTTTGTGAGGTTGACTTAATTTATATAAATATGTTTTTTTGTGTTTTCATGT 1260
1277 TTTAGTTTTTGTGAGTTTAATGACTTATTTATA-----TAAATTTTTTTTGTTCATAT 1332
1261 GAATGAGCGTCTAGCAGAGACTGTGGCCAAAGTCTTTAGTAGCTGTGTTTATCTGTGTAG 1320
1333 TGATGTGTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1392
1321 GACTGTGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380

1393 GACTGTAGAA-----AAGGGAACGAAACATCCAGAGCGGTGTGTGATCAGCTAA 1444
1381 GCTAGCGGTATCCCTCAGCTGTGTGCTGCATA 1411
1445 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1475

RESULT 9
US-08-802-627A-3
; Sequence 3, Application US/08802627A
; Patent No. 5892017
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING PF4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,627A
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 13-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-802-627A-3

Query Match 55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGATTAATCTCTGAGAGAGTGGGCTCTGGAGACTATGACTCCCAACAG 60
Db 106 ATATACACTTCTGATTAATCTCTGAGAGAGTGGGCTCTGGAGACTATGACTCCATGAAG 165
QY 61 GAACCTCTGCTTCGGGATGAAACGTCATTTCAATAGGATCTTCTGCGCCACCATCTAC 120
Db 166 GAACCTCTGCTTCGGTGAAGAAATGCTAATTTCAATAAATCTTCTGCGCCACCATCTAC 225
QY 121 TTCATCATCTTCTGATCTGGCATAGTCGGCATAGGATTCGTGATCTCTGTCATGGGTTAC 180
Db 226 TCCATCATCTTCTTAATCTGTCATTTGTGGCAATGGATTCGTGATCTCTGTCATGGGTTAC 285

181 CAGAGAGAGTAAAGAGCATGACGAGCAAGTACCGCTGACCTGTGCTGAGTGGCTGACCTC 240
Db CAGAGAGAGTAAAGAGCATGACGAGCAAGTACCGCTGACCTGTGCTGAGTGGCTGACCTC 345
241 CTCCTTTGTCATCACTCCCTCTTGGGAGTGTGATGCGATGGCTGATGCTGCTGCTGG 300
Db CTCCTTTGTCATCACTCCCTCTTGGGAGTGTGATGCGGCAAACTGATGCTGCTGG 405
301 AAAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGCAGCTTCTC 360
Db AACTTTCTTATCAAGGAGTCCATGTCTATCTACAGCTCAACCTCTACAGCAGTGTCTC 465
361 ATCTGGCTTCATCAGCTGAGCGGTGATCTGCGCATTTGTCACGCGCAGCAAGTCAA 420
Db ATCTGGCTTCATCAGCTGAGCGGTGATCTGCGCATTTGTCACGCGCAGCAAGTCAA 525
421 AGGCAAGGAACTGCTGGCTGAAAGGAGTCTATGTGGGCGTCTGGATCCCGCCCTC 480
Db AGGCAAGGAGTGTGGCTGAAAGGAGTGTATGTGGGCGTCTGGATCCCGCCCTC 585
481 CTCCTGACTATACCTGATCTCATCTTTGCGAGTCTGAGCGGAGGAGACATGATCAGGG 540
Db CTGCTGACTATTCGCACTTCATCTTTGCCAAGC-----TCAGTGAGGCA 630
541 GATGACAGTATCATCTGAGCGCTTTTACCCGATAGCTGTGGATGGTGTGTTTCAA 600
Db GATGACAGTATCATCTGAGCGCTTTTACCCGATAGCTGTGGATGGTGTGTTTCAA 690
601 TTCCAGCATATAAGTGGGTCTCATCTGCGCGCATGCTGCTCTCTCTCTCTCTCTCT 660
Db TTCCAGCATATAAGTGGGTCTCATCTGCGCGCATGCTGCTCTCTCTCTCTCTCTCT 750
661 ATCATCATCTTAAGTCTGACACTCCAAAGGCGCAGAGAGCGCAAGGCGCTCAAGACG 720
Db ATCATCATCTTAAGTCTGACACTCCAAAGGCGCAGAGAGCGCAAGGCGCTCAAGACG 810
721 ACAGTCACTCTCATCTGAGTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db ACAGTCACTCTCATCTGAGTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
781 ATCGACTCTTCATCTTTGGGAGTCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 840
Db ATCGACTCTTCATCTTTGGGAGTCTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 930
841 CACAAGTGGATCTCCATCAGAGGCGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db CACAAGTGGATCTCCATCAGAGGCGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 990
901 CTCTATGCTCTCTCGGGGCGAAGTTTCAAAAGTCTGCGCGAGGATGCTCAACTCTCATG 960
Db CTCTATGCTCTCTCGGGGCGAAGTTTCAAAAGTCTGCGCGAGGATGCTCAACTCTCATG 1050
961 AGCAGAGGCTCCAGCTCAAGATCTTTTCCAAAGGAGGCGGCTGCTCTCTCTCTCTCTCT 1020
Db AGCAGAGGCTCCAGCTCAAGATCTTTTCCAAAGGAGGCGGCTGCTCTCTCTCTCTCTCT 1110
1021 TCCAGGAGTCAAGATCTCTCAAGTCTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1080
Db TCCAGGAGTCAAGATCTCTCAAGTCTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1162
1081 AAT 1140
Db AAT 1222
1141 ACTGACGAGTCTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1200
Db ACTGACGAGTCTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1276
1201 TCTAGTTTGTGAGGTTGAGTCTTATATATATATATATATATATATATATATATATATAT 1260
Db TCTAGTTTGTGAGGTTGAGTCTTATATATATATATATATATATATATATATATATATAT 1332
1261 GAATGAGCGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320

1333 TGATGTGTGTCTAGCAGGACCTGTGGCAAGTCTTCTAGTTGCTGTATGTCTCTGGTAG 1392
Qy 1321 GACTGTAGAACTGTAGAGGAACTGAACTCCAGAACTGTGTGGTAAATTTGAATAAA 1380
Db 1393 GACTGTAGAA-----AAGGAACTGAACTCCAGAGCGGTGTAGTGAATCAGTAAA 1444
Qy 1381 GCTAGCGTGTATCTCAGCTGTGTCTGCATA 1411
Db 1445 GCTAGAAATGATCCCGCAGCTGTTTATGCATA 1475

RESULT 10

US-08-801-238-3
; Sequence 3, Application US/08801238
; Patent No. 5919896
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PP4A RECEPTOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,238
FILING DATE: 19-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284586
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/076093
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/810782
FILING DATE: 19-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P0706P2P1D1
TELEPHONE: 415/225-5530
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear

US-08-801-238-3

Query Match 55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 18-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACACTTCTGATTAATCTCTGAGAAAGTGGGTCTGGAGACTATGACTCCCAACAAG 60
Db 106 ATATACACTTCTGATTAATCTCTGAGAAAGTGGGTCTGGAGACTATGACTCCCAACAAG 165
Qy 61 GAACCTGTCTCGGAGTGAAGAGTCCATTTCAATAGGATCTTCTGCGCCACCATCTAC 120
Db 166 GAACCTGTCTCGGAGTGAAGAGTCCATTTCAATAGGATCTTCTGCGCCACCATCTAC 225

QY	121	TTCAATCAATCTTTTGA	CTGGCATAGT	CGGC	AAATGGAT	TGGTGAT	CCTGGTCA	TGGGTAC	180
DB	226	TCCATCATCTTTCTTAA	CTGGCAAT	TGTGGGCA	ATGGAT	TGGTCA	TCTCTGGTCA	TGGGTAC	285
QY	181	CAGAAGAAGCTTAAG	CAGCAT	CACGGACA	AGTACCGGCT	GCACCTGT	CAGTGGCT	GCAGCTC	240
DB	286	CAGAAGAAACTGAGA	AGC	ATGACGGA	CAAGTAC	AGGCTGC	ACCTGT	CAGTGGCCGACCTC	345
QY	241	CTCTTTGTCA	TACACT	CTCCCTCT	CTGGCAGT	TGATGCGCA	TGGCTG	ACTGGTACTTTGGG	300
DB	346	CTCTTTGTCA	TACAGCT	CTCCCTCT	CTGGCAGT	TGATGCGG	CAAACTGGT	ACTTTTGGG	405
QY	301	AAATTTTGTGTA	AGGCTGC	CAATAT	CATCTAC	ACTGTCA	ACCTCT	CACAGCAGCGTTCTC	360
DB	406	AAC	TTCTATGCA	AGGCAGT	CCATGT	CATCTAC	ACAGTCA	ACCTCTACAGCAGTGTCTC	465
QY	361	ATCCTGGCCTTCAT	CAGCCTGG	ACGGTAC	CTCCGCA	TGTC	CACGCGCA	CCACAGTCAA	420
DB	466	ATCCTGGCCTTCAT	CAGTCTGG	ACCGGTAC	CTGGCCAT	CGTCC	CACGCGCA	CCACAGTCA	525
QY	421	AGGCAAGAAACTG	CTGGCTG	AAAGG	CAGTCTAT	TGTGGCGG	CTCGAT	TCCCAAGCCCTC	480
DB	526	AGGCCAAGGA	CTGTTGG	CTGTA	AAAGGTGT	CTATGT	TGGCGT	CTCGATCCCTGCCCTC	585
QY	481	CTCTGACTAT	ACTG	CACTTC	ATCTTTG	CCGAGCT	CATCAG	CCAGGGGACATCAGTCA	540
DB	586	CTGCTGACTAT	TC	CCGACTTC	ATCTTTG	CCAAAG	-----	TCAGTGAGGCA	630
QY	541	GATCACAGGTAC	TACTGT	GACGCGCTTT	ATCCCGCAT	AGCCTGT	CGATGGT	TGGTGTTC	600
DB	631	GATGACAGATAT	ATCTGT	GACCGCTT	CTACCCCA	ATGACTT	TGGG	TGGTGTGTTCCAG	690
QY	601	TTCAGCATAT	AAATGGTGG	GTCTCAT	CTCTG	CCCGCAT	CGTCA	TCCTCTCTCTGTTACTGC	660
DB	691	TTTCAGCACAT	CA	TGGTGGCCTT	ATCTCTG	CGCTGTAT	TGTCA	TCTCTCTCTGTTACTGC	750
QY	661	ATCATCATCT	CTAAGCTGT	CACACT	CCAAAGG	CCACAG	AGCGCA	AGGCCCTCAAGACG	720
DB	751	ATTATCATCT	CAAGCTGT	CACACT	CCAAAGG	CCACAG	AGCGCA	AGGCCCTCAAGACG	810
QY	721	ACAGTCA	TCCTCAT	CTAGCTTTCTTT	TG	CCCTGTG	CGTCCAT	ATTATG	780
DB	811	ACAGTCA	TCCTCAT	CTGCTGGCTTTCT	CGCTGT	TGGTCTG	CTTACT	ACA	870
QY	781	ATCCAGCTCCT	CAT	CTTTTGGGAGT	CATCA	AGCAAGAT	GTGACT	TCGAGAGCAT	840
DB	871	ATCCAGCTCCT	CAT	CTCTCTGGA	AAATCAT	CAAGCAG	GGTGTGAGT	TTGAGA	930
QY	841	CACAAGTGGAT	CTCCAT	CACAGAG	CCCTCG	CCCTTCTT	TCCACT	TGTGCTCTGA	900
DB	931	CACAAGTGGAT	TTCCAT	CACAGAG	CCCTAG	CTTTCTT	TCCACT	TGTGCTCTGA	990
QY	901	CTCTATGCTCCT	CTGGGGCC	AGTTCAA	AGCTCT	CGCCAGCA	TGACACT	CACTCCAT	960
DB	991	CTCTATGCTCCT	TTGGAGCC	AAATTTAA	AAACCTCT	GTCCAGCA	CGCAGT	CACTCTGTG	1050
QY	961	AGCAGAGGCT	CCAGCCTCA	AGATCCTTTT	CCAAAGG	AAAGGGGGT	GTGAC	ACTCTTCCGTC	1020
DB	1051	AGCAGAGGCT	CCAGCCTCA	AGATCCTCT	CCAAAGG	AAAGGGGGT	GTGAC	ACTCTTCCGTC	1110
QY	1021	TCCACGAGT	TCAGAA	TCTCCAG	TTTTC	ACTCCAG	CTTAA	CCCTTAT	1080
DB	1111	TCCACTGAGT	CTGAGTCTT	CAAGTTTTC	ACTCCAG	CTTAA	-----	CA	1162
QY	1081	AAATATATATAT	ATATGATA	AAAGAACTTTTT	ATGTTTAC	ACATATTT	TCCAGAT	ATAAGAG	1140
DB	1163	GAC	TTTTTTTATAC	GATAA	TAACTTTTTT	TAAAGTTA	CACATTTT	CAGATATAAAG	1222
QY	1141	ACTGACAG	CTCTGT	TACAGTTTTTTTTTTTT	TTTTTTTT	TAATG	ACTGTT	GGGAGTTAT	1200
DB	1223	ACTGACCAAT	TGTACAG	TTTTTTTAT	TGCTGTT	GGATTTTT	TGT	-----	1276

```

Qy 1201 TCTAGTTTTGTGAGGTTGCACTTAATTTATATAAATATATGTTTTTTGTTGTTTCATGCT 1266
Db 1277 TTTAGTTTTGTGAGGTTTAAATTTGACCTTATTTATA-----TAAATTTTTTTTGTTCATAT 1332
Qy 1261 GAATGACGCTCTAGGCAGGACCTGTGCGCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
Db 1333 TGATGTGTGCTAGGCAGGACCTGTGCGCAAGTTCTTAGTTGCTGTATGCTCTCGTGGTAG 1392
Qy 1321 GACTGTAGAACTGTAGAGGAGAAATGCAATCCAGAAATGTGTGGTAAATTTGAATAAA 1380
Db 1393 GACTGTAGAA-----AAGGGAACCTGCAACATCCAGAGCGGTGTAGTGAATCACGTAAA 1444
Qy 1381 GCTAGCGCTGATCCTCAGCTGTTGCTGCATA 1411
Db 1445 GCTAGAAATGATCCCGACGTTTATGCATA 1475

RESULT 11
US-08-801-228-3
; Sequence 3, Application US/06801228
; Patent No. 5922541
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: METHODS FOR DETECTION AND AMPLIFICATION OF
; TITLE OF INVENTION: PF4A RECEPTOR NUCLEIC ACID
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,228
; FILING DATE: 19-Feb-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284586
; FILING DATE: 10-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2P1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-801-228-3

Query Match 55.1%; Score 933.4; DB 2; Length 1737;
Best Local Similarity 81.8%; Pred. No. 1e-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
Qv 1 ATATACACTTCTGATTAACCTACTCTGAGGAAGTGGGGTCTGGGAGACTATGACTCCAAACAAG 60

```

Db 106 ATATACACTTCACTAATCACTACCGAGAAATGGCTCAGGGACATGACTCCATGAAG 165
Qy 61 GAACCCCTGCTTCGGGATGAAGACGTCATTTCAATAGGATCTTCCTGCCCAACCATCTAC 120
Db 166 GAACCCCTGCTTCGGGATGAAGACGTCATTTCAATAGGATCTTCCTGCCCAACCATCTAC 225
Qy 121 TTCACTATCTTCTGACTGCGCATAGTCGGCAATGGATGGTGATCTGCTGCTGAGGTTAC 180
Db 226 TCCATCATCTTCTTAACCTGCAATGTGGGCAATGGATGGTGATCTGCTGCTGAGGTTAC 285
Qy 181 CAGAAGAGCTAAGAGGACATGACGCAACAGTACCGGCTGACCTGTGACGTGGCTGACCTC 240
Db 286 CAGAAGAACTGAGAAGCATGACGCAACAGTACAGGCTGCACCTGTGAGTGGCGGACCTC 345
Qy 241 CTCCTTCTCATCACACTCCCTTCTGGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 346 CTCCTTCTCATCACACTCCCTTCTGGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG 405
Qy 301 AAATTTTGTGTAAAGCTGCTCATATCATCTACACTGTCAACCTCTACAGCAGCTTCTC 360
Db 406 AACTTCTATGACGAGGAGTCCATGTCTATCACAGTCAACCTCTACAGCAGTGTCTC 465
Qy 361 ATCTGCGCTTCATCAGCTGAGCGGTACCTGCGCAATGCTCAAGCGCAACCAAGTCAA 420
Db 466 ATCTGCGCTTCATCAGCTGAGCGGTACCTGCGCAATGCTCAAGCGCAACCAAGTCAA 525
Qy 421 AGGCAAGGAACTGCTGGCTGAAAGCGAGTCTATGTGGCGCTGCTGCTGCTGCTGCTGCTG 480
Db 526 AGGCAAGGAGCTGCTGGCTGAAAGCGAGTCTATGTGGCGCTGCTGCTGCTGCTGCTGCTG 585
Qy 481 CTCCTGACTATACCTGCTTCTTTCGCGAGCTCAGCCAGGCGGACATCATCAGGCGG 540
Db 586 CTGCTGACTATCCCGACTTCTATCTTTCGCGAG- - - - -TCAGTGAGGCA 630
Qy 541 GATCAGAGTACATCTGTCAGCGCTTACCCGATAGCTGTGGATGGTGGTGGTTCAA 600
Db 631 GATCAGAGTATATCTGTCAGCGCTTACCCGATAGCTGTGGATGGTGGTGGTTCAG 690
Qy 601 TTCAGCATATAATGCTGGTCTCATCTGCGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 691 TTTTCAGCATATGCTGGTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
Qy 661 ATCATCATCTTAAGCTGTACATCTCAAGGCGCACCAAGGCGCAAGGCGGCTCAAGAGC 720
Db 751 ATTATCATCTCAAGCTGTACATCTCAAGGCGCACCAAGGCGCAAGGCGGCTCAAGAGC 810
Qy 721 ACAGTATCTCATCTGCTGCTTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 811 ACAGTATCTCATCTGCTGCTTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
Qy 781 ATCGACTCTTCTATCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 871 ATCGACTCTTCTATCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 930
Qy 841 CACAAGTGGATCTCCATCAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 931 CACAAGTGGATCTCCATCAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 990
Qy 901 CTCCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 991 CTCCTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
Qy 961 AGCAGAGGCTCCAGCTCAAGATCTTTCAGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 1051 AGCAGAGGCTCCAGCTCAAGATCTTTCAGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCT 1110
Qy 1021 TCCACGAGTCAAGATCTTTCAGAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1111 TCCACTGAGTCTGAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 1162
Qy 1081 AAT 1140

RESULT 12

US-09-104-296-3
; Sequence 3, Application US/09104296
; Patent No. 6087475
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PPA4 Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinFatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,296
; FILING DATE: 24-June-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/664228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single

TOPOLOGY: Linear
US-09-104-296-3
Query Match 55.1%; Score 933.4; DB 3; Length 1737;
Best Local Similarity 81.8%; Pred. No. 18-210;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
1 ATATACACTTTCGATTAACACTCTGAGAAAGTGGGGTCTGGAGACTATGACTCCAAACAG 60
106 ATATACACTTTCGATTAACACTCTGAGAAAGTGGGGTCTGGAGACTATGACTCCAAACAG 165
61 GAACCCCTGCTCCGGGATGAAGAGTGCATTTCAATAGGATCTTCTGCGCCACCACTAC 120
166 GAACCCCTGCTCCGGGATGAAGAGTGCATTTCAATAGGATCTTCTGCGCCACCACTAC 225
121 TTATCATCTCTCTTGTGCTGCTAGTGCATTCGCAATGCAATGCAATGCAATGCAATGCAAT 180
226 TCATCATCTCTCTTGTGCTGCTAGTGCATTCGCAATGCAATGCAATGCAATGCAATGCAAT 285
181 CAGAAGAGCTAAGGAGCATGACGCAAGATGACGCGCTGCACTGTGCTGCTGCTGCTGCTGCT 240
286 CAGAAGAGCTAAGGAGCATGACGCAAGATGACGCGCTGCACTGTGCTGCTGCTGCTGCTGCT 345
241 CTCTTTGTCTATCAGCTCCCTCTCTGCGAGTGTGATGCGCATGCTGCTGCTGCTGCTGCTGCT 300
346 CTCTTTGTCTATCAGCTCCCTCTCTGCGAGTGTGATGCGCATGCTGCTGCTGCTGCTGCTGCT 405
301 AAATTTTGTGTAAGGCTGTCTATATCATCTACACTGTCAACTGTCAACTGTCAACTGTCAACT 360
406 AACTTCTATGCAAGGAGTCTGCTATCTATCACTCACTCACTCACTCACTCACTCACTCACTCA 465
361 ATCTGCGCTTCACTGAGCTGAGCGTACTGCGCATGCTGCGCATGCTGCGCATGCTGCGCATGCT 420
466 ATCTGCGCTTCACTGAGCTGAGCGTACTGCGCATGCTGCGCATGCTGCGCATGCTGCGCATGCT 525
421 AGGCAAGGAACTGCTGGCTGAAAGGAGTCTATGCTGGGCGTCTGGATGCCAGCCCTC 480
526 AGGCAAGGAACTGCTGGCTGAAAGGAGTCTATGCTGGGCGTCTGGATGCCAGCCCTC 585
481 CTCCTGATATACCTGATCTTCTTGGCGAGTCTGCGCATGCTGCGCATGCTGCGCATGCTGCG 540
586 CTCCTGATATACCTGATCTTCTTGGCGAGTCTGCGCATGCTGCGCATGCTGCGCATGCTGCG 630
541 GATCAGAGTACATCTGAGCGCTTCTTACCGCATGCTGCGCATGCTGCGCATGCTGCGCATGCT 600
631 GATCAGAGTACATCTGAGCGCTTCTTACCGCATGCTGCGCATGCTGCGCATGCTGCGCATGCT 690
601 TTCCAGCATATAATGTTGGGCTCTCATCTGCGCGCATGCTGCGCATGCTGCGCATGCTGCGCAT 660
691 TTTCAGCATATAATGTTGGGCTCTCATCTGCGCGCATGCTGCGCATGCTGCGCATGCTGCGCAT 750
661 ATCATCTCTTACCTGATCTACACTCCAGGCGCCACCAAGGCGCAAGGCGCCCTCAAGAG 720
751 ATTATCATCTCAAGCTGTACACTTCCAGGCGCCACCAAGGCGCCACCAAGGCGCCCTCAAGAG 810
721 ACAGTCACTCTCATCTGATCTTCTTGGCGAGTCTGCGCATGCTGCGCATGCTGCGCATGCTGCG 780
811 ACAGTCACTCTCATCTGATCTTCTTGGCGAGTCTGCGCATGCTGCGCATGCTGCGCATGCTGCG 870
781 ATGAGCTCTCTATCTTGGGAGTCTATCAAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
871 ATGAGCTCTCTATCTTGGGAGTCTATCAAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
841 CACAAGTGGATCTCCATCAGAGGCGCCCTGCGCTTCTTCCACTGTGCTGCTGCTGCTGCTGCT 900
931 CACAAGTGGATCTCCATCAGAGGCGCCCTGCGCTTCTTCCACTGTGCTGCTGCTGCTGCTGCT 990
901 CTCTATGCTTCTCTGCGGCGCAAGTTCAAAAGCTCTGCGGAGCATGCTGCTGCTGCTGCTGCTGCT 960
991 CTCTATGCTTCTCTTGGAGCGCAATTTAAACCTCTGCGGAGCGCACTGCTGCTGCTGCTGCTGCT 1050
961 AGCAGGCTCTCAGCTCAAGATCTTCTTCCAGGAAAGCGGGTGGACACTCTTCTCGCTC 1020

Db 1051 AGCAGAGGCTCCAGCTCAAGATCTCTCCAAAGGAAGCGAGGTGGACATTCATCTGTT 1110
QY 1021 TCCACGAGTCCAGATCTCCAGTCTTCTCACTCCAGCTAACCCCTTATGCAAGACCTTATAT 1080
Db 1111 TCCACTGAGTCTGAGTCTTCAAGTCTTCTCACTCCAGCTAA-----CACAGATGTAATA 1162
QY 1081 AAT 1140
Db 1163 GACTTTTTTATACGATAAATAAATCTTTTTTATAGTTACACATTTTTCAGATATAAAG 1222
QY 1141 ACTGACAGCTCTTGTACAGTTTTTTTTTTTTTTTTTAAATGACTGTGAGGATTTATGTTCC 1200
Db 1223 ACTGACCAATATGTACAGTTTTTATGCTTGTGATTTTCT-----CTTGTGTTTC 1276
QY 1201 TCTAGTTTTTGTGAGGTTTCTGACTTAATATATATATATATATATATATATATATATATAT 1260
Db 1277 TTTAGTTTTTGTGAGGTTTAAATGACTTATATATA-----TAAATTTTTTGTGTTTCA 1332
QY 1261 GAATGAGCGTCTAGCGAGGACCTGTGCGCAAGTTCTTTAGTAGCTGTTTATCTGTTGTAG 1320
Db 1333 TGATGTGTCTAGCGAGGACCTGTGCGCAAGTTCTTTAGTTGCTGTATGCTCTCGTGTAG 1392
QY 1321 GACTGTGAGACTGTAGAGGAAGAACTGAACATTCAGAAATGCTGCTGCTGCTGCTGCTGCT 1380
Db 1393 GACTGTAGAA-----AAGGAACTGAACATTCAGAGCGTGTAGTAATCAGGTAA 1444
QY 1381 GCTAGCGTGTATCTCAGCTGTTGCTGCATA 1411
Db 1445 GCTAGAAATGATCCCGAGCTGTTATGCATA 1475
RESULT 13
PCT-US94-06380-2
; Sequence 2, Application PC/TUS9406380
; GENERAL INFORMATION:
; APPLICANT: Chunharapai, Anan
; APPLICANT: Lee, James
; APPLICANT: Hebert, Caroline
; APPLICANT: K. Jin Kim
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY DISORDERS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06380
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: 706P2P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 bases
; TYPE: nucleic acid

QY	1	ATATACACTTCTGTGATACTACTCTGAGAGATGGGCTCTGGAGACTATGACTTCCAACAAG	60
DB	41	ATATACACTTCTGAGATAACTACACGAGAAATGGGCTCAGGGACATGATGATCTCCATGAAG	100
QY	61	GAACCCCTGCTTCCGGGATGAAGCGTCATTTCAAATAGGATCTTCTGCGCCACCACTAC	120
DB	101	GAACCCCTGTTTCCGTTGAGAAATGCTAAATTTCAATAAAATCTTCTGCGCCACCACTAC	160
QY	121	TTTCATCACTTCTTGGACTGTCGNATAGTCGGCAATGGATGGTGATCTCTGTCATGAGGTTAC	180
DB	161	TCCATCATCTTCTTAACTGGCATTTGTGGGCATTTGGATGGTCACTCTGTCATGAGGTTAC	220
QY	181	CAGAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCAGTGGCTGACCTC	240
DB	221	CAGAAGAAGCTAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCGCCACCTC	280
QY	241	CTCTTTGTTCATCACACTCCCTCTCTGGGCAAGTATGGATGGCATGGCTGACTGTTCTTGGG	300
DB	281	CTCTTTGTTCATCACGCTTCCCTCTCTGGGCAAGTATGGATGGCATGGCTGACTGTTCTTGGG	340
QY	301	AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACTGTACAAGAGCGTTCTC	360
DB	341	AACTTCTTATGCAAGGCAAGTCCATGTCACTACACAGTCAACCTGTACAGCAGTGTCCCTC	400
QY	361	ATCTCGGCTTCATCAGCTGGACCGGTACTCGCCATTGTCCAGGCCACCAAGTCAAA	420
DB	401	ATCTCGGCTTCATCAGCTGGACCGGTACTCGCCATTGTCCAGGCCACCAAGTCAAA	460
QY	421	AGGCCAAGGAACTCTGGCTGTAAGGCAAGTCTATGTGGCGCTCTGGATCCCAAGCCCTC	480
DB	461	AGGCCAAGGAACTCTGGCTGTAAGGCAAGTCTATGTGGCGCTCTGGATCCCAAGCCCTC	520
QY	481	CTCTCGACTATACCTGACTTCACTTTTGGCGAGTCTCAGCCAGGGGACATCAGTCAAGGG	540
DB	521	CTGCTGACTATTCGCGACTTCACTTTTGGCACAAG-----TCACTGAGGCA	565
QY	541	GATGACAGGTACATCTGTACCGCTTTACCCCGATAGCCTGTGGATGATGCTTTCAA	600
DB	566	GATGACAGATATATCTGTGACCGCTTACCCCAATGACTTTGTGGGTGGTGTGTTCCAG	625
QY	601	TTTCCAGCATATAAGTGGTGGGTCTCATCTCCCGCGATCTCATCTCTCTCTGTTACTGCG	660
DB	626	TTTCCAGCATATGAGTGGTGGGCTTATCTGCTCTGTATTGTCACTCTCTCTCTGTTACTGCG	685
QY	661	ATCATCATCTCTTAAGCTGTCACTTCCAAGGGCCACCAAGAGCGCGAAGGCCCTCAAGACG	720
DB	686	ATTATCATCTCCAAGCTGTCACTTCCAAGGGCCACCAAGAGCGCGAAGGCCCTCAAGACG	745
QY	721	ACAGTCATCTCATCTGACTTTCTTTGCTGCTGGCTGCCATTAATTATGTGGGATCAGC	780
DB	746	ACAGTCATCTCATCTGACTTTCTTTGCTGCTGGCTGCCATTAATTATGTGGGATCAGC	805
QY	781	ATCGACTCTCTCATCTTTTGGGAGTCACTCAAGCAAGGATGTGACTTTCAGAGCATTTGTG	840
DB	806	ATCGACTCTCTCATCTCTCGAAATCATCAAGCAAGGATGTGAGTTTGAAGCACTGTG	865
QY	841	CACAGTGTGATCTCCATCACAGAGCCCTGCGCTTCTTCCACTGTGTGCTGAGACCCCATC	900
DB	866	CACAGTGTGATTTCCATCACAGAGCCCTGAGCTTCTTCCACTGTGTCTGAGACCCCATC	925
QY	901	CTCTATGCTCTCTCTGGGCGCAAGTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATG	960

CLASSIFICATION: 514
PRIOR APPLICATION DATA: US 07/977,452
APPLICATION NUMBER: 17-NOV-1992
FILING DATE: 17-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 5759804and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/POCKET NUMBER: 31794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1317 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 201..1211
US-08-153-848-45

US-08-153-848-45

Query Match 49.2%; Score 834; DB 1; Length 1317;
Best Local Similarity 84.1%; Pred. No. 2.5e-187;
Matches 974; Conservative 0; Mismatches 160; Indels 24; Gaps 2;

QY 1 ATATACACTTCTGTAATACTACTCTGAAGAAGTGGGTCTCGAGACTATGACTCCACAAG 60
Db 171 ATATACACTTCTGTAATACTACTCACGAGAAATGGCTCAGGGACTATGACTCCAGAAG 230
QY 61 GAACCTCTGCTTCCGGGATGAACAGTCCATTTCAATAGGATCTTCTGCCCCACCATCTAC 120
Db 231 GAACCTCTGCTTCCGTGAAGAATAATGCTAAATTCATTAATAATCTTCTGCCCCACCATCTAC 290
QY 121 TTATCATCTTCTTGATCTGGCATAGTCTGGCAATGAGATGTGATCTCTGCTGATGGGTAC 180
Db 291 TCCATCATCTTCTTAATCTGGCAATGTGGCAATGAGATGTGATCTCTGCTGATGGGTAC 350
QY 181 CAGAAGAAGCTAAGCAGCATGACGACCAAGTACCGCTGCACCTGTCTAGTGGCTGACCTC 240
Db 351 CAGAAGAAGCTAAGCAGCATGACGACCAAGTACCGCTGCACCTGTCTAGTGGCTGACCTC 410
QY 241 CTCCTTTGTCTACACTCCCTCTCTGGGCAAGTGTGATGCCATGCTGACTGCTACTTTGGG 300
Db 411 CTCCTTTGTCTACACTCCCTCTCTGGGCAAGTGTGATGCCATGCTGACTGCTACTTTGGG 470
QY 301 AAATTTTGTGTAGGCTGCCATATCATCTACATGCTCAACCTCTACAGCAGCGTTCTC 360
Db 471 AACTTCTATGCAAGGAGTCCATGTCTATACACAGTCAACCTCTACAGCAGTGCCTC 530
QY 361 ATCTGCGCTTCAATGACCTGACCGGTACTCTGCCATTTGCCAGCCCAACAGTCAA 420
Db 531 ATCTGCGCTTCAATGACCTGACCGGTACTCTGCCATTCGTCAGCCCAACAGTCAAG 590
QY 421 AGGCCAGGAACCTGCTGGAAGCGAGTCTATGTGGGGCTCTGGATCCCGAGCCCTC 480
Db 591 AGGCCAGGAAGCTGTTGGGTGAAAGGTGTATGTTGGGGTCTGGATCCCGAGCCCTC 650
QY 481 CTCCTGACTATACCTGACTTTCATCTTTGGCGAAGTCTAGCCAGGGGGACATCAGTCAGGGG 540
Db 651 CTGCTGACTATTCGGAATTCATCTTTGCCAAG-----TCAGTGAAGCA 695
QY 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTTTCAA 600
Db 696 GATGACAGATATATCTGTGACCGCTTTACCCGATAGCTGTGGTGGTGGTGGTTCAG 755
QY 601 TTCCAGCATATAATGGTGGGTCTCATCTGCGGGCATCTCATCTCTCTCTCTTACTGC 660
Db 756 TTTCCAGCATATGTTGGCTTATCTGCTGGTATTTGTCATCTCTCTCTCTTATTC 815
QY 661 ATCATCATCTTAAGTGTGACACTCCAGGGGCCACAGAGGCCAAGGCCCTCAAGAGC 720
Db 816 ATTATCATCTCAAGCTGTACACTCCAGGGGCCACAGAGGCCAAGGCCCTCAAGAGC 875
QY 721 ACAGTCACTCTCATCTCTGCTTCTTCTGCTGCTGCTGCTATATATGTTGGGGATCAGC 780
Db 876 ACAGTCACTCTCATCTCTGCTTCTTCTGCTGCTGCTGCTTCTGCTGCTTCTGCTGCTGCT 935
QY 781 ATCGACTCTCTCATCTCTTGGGAGTCTCAAGCAGGATGTGACTTTCCAGAGCATTTGTG 840
Db 936 ATCGACTCTCTCATCTCTTGGGAGTCTCAAGCAGGATGTGACTTTCCAGAGCATTTGTG 995
QY 841 CACAAGTGGATCTCCATCAGAGGCGCTCGCTTCTTCCACTGTTGCTGGAACCCCATC 900
Db 996 CACAAGTGGATTTCCATCAGAGGCGCTTAGCTTTCTTCCACTGTTGCTGGAACCCCATC 1055
QY 901 CTCATGCTCTCTCGGGGCCAAGTTCAAAAGTCTGCCCCAGCATGCACTCAACTCCATG 960
Db 1056 CTCATGCTCTCTCGGGGCCAAGTTCAAAAGTCTGCCCCAGCATGCACTCAACTCCATG 1115
QY 961 AGCAGAGGCTCCAGGCTCAAGATCTTTTCCAAAGGAAGGGGGTGGACACTCTTCCGTC 1020
Db 1116 AGCAGAGGCTCCAGGCTCAAGATCTCTCCAAAGGAAGGGGGTGGACACTCTTCCGTC 1175
QY 1021 TCCACGGAGTCAGAAATCCTCCAGTTTTCATCTCCAGCTAACCCCTATGTCAAAGACTTATAT 1080

Db 1176 TCCACTGAGTCTGAGTCTTCAAGTTTTTCACTCCAGCTAACACAGATGTAAA----- 1226
QY 1081 AATATATATATATATATGATAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG 1140
Db 1227 AGACTTTTTTATACGATAAATAACCTTTTTTAAAGTTACACATTTTTCAGATATAAAG 1286
QY 1141 ACTGACCAAGTCTTGTACA 1158
Db 1287 ACTGACCAATATTGAAAA 1304

Search completed: May 17, 2004, 16:15:55
Job time : 145.63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2004, 12:11:00 ; Search time 829.19 Seconds
(without alignments)
9271.041 Million cell updates/sec

Title: US-09-367-052-7

Perfect score: 1694
Sequence: 1 atatacacttctgataacta.....taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2947324 seqs, 2269024515 residues

Total number of hits satisfying chosen parameters: 5894648

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	933.8	55.1	5161	16	US-10-160-401-1
2	933.4	55.1	1664	17	Sequence 1, Appl1
3	933.4	55.1	1679	13	Sequence 1213, Ap
4	933.4	55.1	1679	13	Sequence 80, Appl
5	933.4	55.1	1679	13	Sequence 9, Appl1
6	933.4	55.1	1679	13	Sequence 912, Appl
7	933.4	55.1	1679	13	Sequence 912, Appl
8	933.4	55.1	1679	13	Sequence 332, Appl
9	933.4	55.1	1679	15	Sequence 75, Appl
10	933.4	55.1	1679	15	Sequence 58, Appl
11	933.4	55.1	1679	16	Sequence 177, Appl
12	933.4	55.1	1711	10	Sequence 20, Appl
13	933.4	55.1	1711	15	Sequence 459, Appl
14	933.4	55.1	1737	9	Sequence 3, Appl1

```
15 933.4 55.1 1737 13 US-10-666-689-3
16 931.4 55.0 1670 9 US-09-880-107-2143
17 931.4 55.0 1670 10 US-09-960-706-636
18 931.4 55.0 1670 16 US-10-372-683-1
19 931.4 55.0 1670 16 US-10-440-464-77
20 855.6 50.5 1225 15 US-10-101-510-741
21 855.6 50.5 1225 15 US-10-305-720-1235
22 843.6 49.8 1224 15 US-10-101-510-674
23 806.2 47.6 1059 13 US-09-813-651B-84
24 806.2 47.6 1059 15 US-10-014-322A-125
25 806.2 47.6 1059 16 US-10-160-401-2
26 806.2 47.6 1059 9 US-09-870-759-143
27 804.6 47.5 1102 10 US-09-751-708A-143
28 801.4 47.3 1059 11 US-09-826-509-484
29 553.6 32.7 1902 9 US-09-953-692-1
30 553.6 32.7 1902 9 US-09-953-717-1
31 530 31.3 1076 15 US-10-017-161-789
32 530 31.3 1076 16 US-10-292-798-687
33 345.2 20.4 507 13 US-10-085-783A-14955
34 345.2 20.4 507 16 US-10-242-535A-14955
35 303.2 17.9 421 9 US-09-796-692-2611
36 303.2 17.9 421 15 US-10-040-862-2611
37 303.2 17.9 421 16 US-10-057-475B-2611
38 303.2 17.9 421 16 US-10-154-884B-2611
39 287 16.9 487 9 US-09-796-692-8936
40 287 16.9 487 15 US-10-040-862-8936
41 287 16.9 487 16 US-10-057-475B-8936
42 287 16.9 487 16 US-10-154-884B-8936
43 198.4 11.7 1119 12 US-09-997-722-162
44 198.4 11.7 1119 15 US-10-251-385-65
45 198.4 11.7 1679 9 US-09-104-063-5
```

ALIGNMENTS

```
RESULT 1
US-10-160-401-1
; Sequence 1, Application US/10160401
; Publication No. US20030207281A1
; GENERAL INFORMATION:
; APPLICANT: Genesance Pharmaceuticals, Inc.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Bielecki, Karyn M.
; APPLICANT: Koshi, Beena
; APPLICANT: Monroe, Eileen
; APPLICANT: Rounds, Glenn
; TITLE OF INVENTION: HAPLOTYPES OF THE CXCR4 GENE
; FILE REFERENCE: MMH-0121US
; CURRENT APPLICATION NUMBER: US/10/160,401
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: PCT/US01/12268
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,025
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3118)..(3118)
; OTHER INFORMATION: PSI: polymorphic base thymine or adenine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3583)..(3583)
; OTHER INFORMATION: PS2: polymorphic base cytosine or thymine
; FEATURE:
; NAME/KEY: allele
; LOCATION: (3952)..(3952)
; OTHER INFORMATION: PS3: polymorphic base cytosine or thymine
```

901	QY	CTCTATGCTTCTCGGGCCCAAGTTCCTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	960
4070	DB	CTCTATGCTTCTCGGGCCCAAGTTCCTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	4129
961	QY	AGCAGAGCTCCAGGCTTCAAGATCTCTTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	1020
4130	DB	AGCAGAGCTCCAGGCTTCAAGATCTCTTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	4189
1021	QY	TCCAGGAGTCAGATCTCCAGTTCCTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	1080
4190	DB	TCCAGGAGTCAGATCTCCAGTTCCTCCAAAGGAAACGCGGGTGGGACACTCTTCCCGTC	4241
1081	QY	AAAT	1140
4242	DB	GACTTTTTTTTATACGATAAATAAATTAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT	4301
1141	QY	ACTGACGAGCTCTTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	1200
4302	DB	ACTGACGAGCTCTTGACAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT	4355
1201	QY	TCTAGTTTTTGTGAGGTTTGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1260
4356	DB	TCTAGTTTTTGTGAGGTTTGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	4411
1261	QY	GAATGAGGCTCTGACGAGGCTGTGGGCAAGTCTTAGTAGCTCTTAGTAGCTCTTAGTAGCTCT	1320
4412	DB	GAATGAGGCTCTGACGAGGCTGTGGGCAAGTCTTAGTAGCTCTTAGTAGCTCTTAGTAGCTCT	4471
1321	QY	GACTGTAGAAGCTGTAGAGGAAAGAACTGAACATTCAGAAATGTGGTAAATTTGAATAAA	1380
4472	DB	GACTGTAGAAGCTGTAGAGGAAAGAACTGAACATTCAGAAATGTGGTAAATTTGAATAAA	4523
1381	QY	GCTAGCGGTGATCTCAGCTGTGCTGCATA	1411
4524	DB	GCTAGCGGTGATCTCAGCTGTGCTGCATA	4554

RESULT 2

US-10-641-643-1213

Sequence 1213, Application US/10641643

Publication No. US20040077003A1

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

Susan G. Stuart

Jeffrey J. Seilbamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

ue May 18 13:54:46 2004

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1213:
SEQUENCE CHARACTERISTICS:
LENGTH: 1664 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g219868
SEQUENCE DESCRIPTION: SEQ ID NO: 1213 :

US-10-641-1213
Query Match 55.1%; Score 933.4; DB 17; Length 1664;
Best Local Similarity 81.8%; Pred. No. 2.1e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
1 ATATACACTCTCTGATACTACTCTGAAGAAGTGGGCTGGAGACTATGACTCCCAACAAG 60
2 ATATACACTCTGATACTACTCAACGGAAGTGGGCTCAGGGGACTATGACTCCATGAAG 151
61 GAACCTCTCTCCGGAGTGAAGAAAGTCCATTTCAATAGGATCTTCCTGGCCCAACATCTAC 120
152 GAACCTCTCTCCGGAGTGAAGAAAGTCTAAATTTCAATAAATCTTCCTGGCCCAACATCTAC 211
121 TTCAATCATCTCTGACATGCGCATAGTCGCAATGATGGTGGTATCCTGGTCAATGGGTTAC 180
212 TCCATCATCTCTTAACTGGCAATTTGGGCAATGATTTGGTCAATCTCTGGTCAATGGGTTAC 271
181 CAGAAGAAAGCTAAGAGGACATGACGACAAAGTACCGGCTGCACCTGTGCATGGCTGACCTC 240
272 CAGAAGAAAGCTAAGAGGACATGACGACAAAGTACAGGCTGCACCTGTGCATGGGCACTC 331
241 CTCCTTTGTCATCAGTCCCTCTTGGGAGTGTGATGCGATGGTGTGATGCTGCTTTGGG 300
332 CTCCTTTGTCATCAGTCCCTCTTGGGAGTGTGATGCGGCAATGCTGATCTTTGGG 391
301 AAATTTTGTGTAAGGCTGTCCATATCATCTACATCTGCAACCTCTACAGCAGGCTTCTC 360
392 AATTTTCTATGCAAGGAGTCAATGCTATCATACAGTCAACCTCTACAGCAGTGTCTC 451
361 ATCTGGGCTTCATCAGCCTGACCGGTACCTCGGCATGTCCAGCCCAACCAAGTCAA 420
452 ATCTGGGCTTCATCAGTCTGACCGGTACCTCGGCATGTCCAGCCCAACCAAGTCAA 511
421 AGGCCAAGAAAGCTGCTGCTGAAAGGAGTCTATGTGGGCTGTGGATCCGACGCTC 480
512 AGGCCAAGAAAGCTGCTGCTGAAAGGAGTGTGATGTGGGCTGTGGATCCGACGCTC 571
481 CTCCTGACTATACCTGACTTCACTTTGCCAGCTGTCAGCCAGGGGACATCAGTCAGGGG 540
572 CTCCTGACTATACCTGACTTCACTTTGCCAGCTGTCAGCCAGGGGACATCAGTCAGGGG 616
541 GATGACAGGTATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGTGTTCAA 600
617 GATGACAGGTATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGTGTTCAA 676
601 TTCCAGCATATATGATGGTGGTCTCATCTCTGCCGCTGCTGATCTCTCTGTTTACTGC 660
677 TTCCAGCATATGATGGTGGTCTCATCTCTGCCGCTGCTGATCTCTCTGTTTACTGC 736
661 ATCATCATCTTAAGTGTGACACTCCAGGCGCAACCAAGCGCAAGCGCCCTCAAGAGC 720
737 ATATCATCTCAAGTGTGACACTCCAGGCGCAACCAAGCGCCCTCAAGAGC 796
721 ACAGTCACTCTCATCTGACTTTCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
797 ACAGTCACTCTCATCTGACTTTCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
781 ATCGACTCTCTCATCTTTTGGAGTCAATCAAGCAAGGATGTGACTTTCAGAGCAATTTGT 840
857 ATCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGATGTGACTTTCAGAGCAATTTGT 916

841 CACAAGTGGATCTCCATCACAAGAGCCCTCGCTTCTTCCACTGTGTGCTGAACCCCATC 900
917 CACAAGTGGATTTCCATCACCAGGCCCTAGCTTCTTCCACTGTGTGCTGAACCCCATC 976
901 CTCTATGCTTCTCTCGGGCCCAAGTTCAAAAGCTCTGCCAGACTGCACTCAACTCCATG 960
977 CTCTATGCTTCTCTTGGAGCCAAATTTAAAGCTCTGCCAGAGCACTCACTCTGTG 1036
961 AGCAGAGGCTCCAGGCTCAAGATCTTCCAAAGAAAGAGCGGGGGGACACTCTTCCGTC 1020
1037 AGCAGAGGCTCCAGGCTCAAGATCTCTCCAAAGAAAGAGCGGGGGGACACTCTTCTGT 1096
1021 TCCAGGAGTCAAGATCTCTCCAGTCTTCACTCCAGCTAAACCTTATGCAAGAACTATAT 1080
1097 TCCAGTGTCTGAGTCTTCAAGTCTTCACTCCAGCTAAACCTTATGCAAGAACTATAT 1148
1081 AAT 1140
1149 GACTTTTTTTTATACGATAAATAAATCTTTTTTATAGTTACACATTTTTTCAATATAAAG 1208
1141 ACTGACCACTCTGTACAGTCTTTTTTTTTTTTTTTTTTTTTTAAATGACTGTGGGAGTTTATGTTCC 1200
1209 ACTGACCAATATGTACAGTCTTTTTTATGCTTGTGATTTTTGT-----CTTGTGTTTC 1262
1201 TCTAGTTTTGTGAGGTTGACTTAATTTATATATATATATATATATATATATATATATATAT 1260
1263 TTTAGTTTTTGTGAACTTTAATTTGACTTTATTTATA-----TAAATTTTTTTTGTTCATAT 1318
1261 GATGAGCTCTAGGAGGACCTGTGGCAAGTCTTCTAGTGTGTATCTGTGTGTAG 1320
1319 TGTGTGTGTCTAGGAGGACCTGTGGCAAGTCTTCTAGTGTGTATCTGTGTGTAG 1378
1321 GACTGTAGAACTGTAGAGGAAAGAACTGAACTTCAGAAATGTGTGTGTAAATGAATAAA 1380
1379 GACTGTAGAA-----AAGGAACTGAACTTCAGAGCGGTGTAGTTAATCAGCTAAA 1430
1381 GCTAGCGGTGATCTCCTCAGCTGTGCTGCATA 1411
1431 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1461

RESULT 3
US-10-211-462-80
Sequence 80, Application US/10211462
Publication No. US20040033495A1
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Aziz, Natsheh
TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
FILE REFERENCE: 018501-006200US
CURRENT APPLICATION NUMBER: US/10/211,462
PRIOR FILING DATE: 2003-02-13
PRIOR APPLICATION NUMBER: US 09/784,356
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 09/791,390
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: US 60/310,025
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: US 60/334,244
PRIOR FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 230
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 80
LENGTH: 1679
TYPE: DNA
ORGANISM: Homo sapiens
US-10-211-462-80

Db	1109	TCAC	TGAGCTGAGGCTTCAAGTTTTCACTCAGCTAA-----CACAGATGTAAA	1160
QY	1081	AA	TATATATATATATGATAAAGAACCTTTTTATGTTACACATTTTCCAGATATAAGAG	1140
Db	1161	GACT	TTTTTTTATACGATAAATAAAGTTTTTTTAAAGTTACACATTTTTCAGATATAAAG	1220
QY	1141	ACTG	ACCAGCTGTGTACAGTTTTTTTTTTTTTTTAAATGACTGCTGGGAGTTTATGTTCC	1200
Db	1221	ACTG	ACCAATATGTACAGTTTTTATGCTGTGGATTTTGT-----CTTGTGTTTC	1274
QY	1201	CT	AGTTTTTGTGAGGTTTGACTTAATTTATATAAATATTTTGTGTTGTTCAATGT	1260
Db	1275	TT	TAGTTTTTGTGAAGTTTAAATTGACTTATTTATA---TAAATTTTTTTTGTTCATAT	1330
QY	1261	GA	TGAGCGTCTAGGACGACCTGTGGCCAAAGTTCTTTAGTAGTGTTTATCTGTGTGTAG	1320
Db	1331	TG	ATGTGTCTAGGACGACCTGTGGCCAAAGTTCTTTAGTTGCTGTATGCTCGTGTGTAG	1390
QY	1321	GACT	GTAGAACTGTAGAGGAAGAACTGAACATTCAGAAATGTGTGTTAAATTCAAATAAA	1380
Db	1391	GACT	GTAGAA-----AAGGAACTGAACATTCAGAGCGGTAGTGAATCACGTAAA	1442
QY	1381	GCT	AGCCGTGATCCTCAGCTGTGTGCATA	1411
Db	1443	GCT	AGAAATGATCCCGAGCTGTTTATGCATA	1473

RESULT 4

US-10-181-906-9

Sequence 9, Application US/10181906

Publication No. US20040053864A1

GENERAL INFORMATION:

APPLICANT: Karsenty, Gerard

APPLICANT: Amling, Michael

APPLICANT: Ducey, Patricia

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROL OF BONE FORMATION VIA

TITLE OF INVENTION: MODULATION OF NEUROPEPTIDE Y ACTIVITY

FILE REFERENCE: 9142-020-999

CURRENT APPLICATION NUMBER: US/10/181,906

PRIOR FILING DATE: 2002-11-13

PRIOR APPLICATION NUMBER: PCT/US01/02040

PRIOR FILING DATE: 2001-01-22

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patent in version 3.0

SEQ ID NO 9

LENGTH: 1679

TYPE: DNA

ORGANISM: Homo sapiens

US-10-181-906-9

Query Match 55.1%; Score 933.4; DB 13; Length 1679;

Best Local Similarity 81.8%; Pred. No. 2.1e-236;

Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY	1	AT	ATACACTCTCTGATAACTACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCCAACAAG	60
Db	104	AT	ATACACTTTCATTAACATACACCGAGGAATGGGCTCAGGGACTATGACTCCATGAAG	163
QY	61	GA	ACCCCTGCTTCGGGATGAAAAACGTCCTCAATAGAGTCTTCTCGCCCAACCATCTAC	120
Db	164	GA	ACCCCTGTTCCGTGAAGAAATAGTCAATTTCAATAAAATCTTCTCGCCCAACCATCTAC	223
QY	121	TT	CATCATCTCTTGTAGTCGCATAGTCGGCAATGGAATGTGTGATCTCTGTGATGGGTTAC	180
Db	224	TC	CATCATCTCTTAACTGTGCATTTGGGCAATGGATTTGCTCATCTCTGTGATGGGTTAC	283
QY	181	CAGA	GAAGCTAAGGAGCATGACGCAACAGTACCGGCTGCACCTGTGCAGTGGGTGACCTC	240
Db	284	CAGA	GAAGCTGAGAGCATGACGCAACAGTACCGGCTGCACCTGTGCAGTGGGTGACCTC	343
QY	241	CT	TTTGTGCATCACCTCCCTTTCTGGGCAAGTTGATGCCCATGGCTGATCTGTTGGG	300

Db 344 CTCTTTGTCATCAGCTTCCCTTCTGGCAGTGTGATGCGGTGCAAACTGGTACTTTGGG 403
Qy 301 AAATTTTGTGTAAGGCTGTCCATATCATCTACATCTGTCACCTCTACAGCAGCTTCTC 360
Db 404 AACTTCTATGCAAGGAGTCCATGTCTATCTACACAGTCAACCTCTACAGCAGTGTCTC 463
Qy 361 ATCTGCGCTTCTATCAGCTGAGCCGCTACCTCGCATTTGTCACCGCCACCAACAGTCAA 420
Db 464 ATCTGCGCTTCTATCAGTCTGAGCAGCTCTCTGCGCATCTGTCACCGCCACCAACAGTCAA 523
Qy 421 AGGCCAAGAACTGCTGCTGAAAGGCAAGTCTATGTTGGCGCTCTGATCCCGCCCTC 480
Db 524 AGGCCAAGAACTGTTGGCTGAAAGGCTCTATGTTGGCGCTCTGATCCCGCCCTC 583
Qy 481 CTCTGACTATCTGACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 584 CTCTGACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
Qy 541 GATGACAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 629 GATGACAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688
Qy 601 TTCAGCATATATGTTGGTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 689 TTTGAGCAGATCATGTTGGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Qy 661 ATCATCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 720
Db 749 ATTATCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 808
Qy 721 ACAGTCATCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 780
Db 809 ACAGTCATCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 868
Qy 781 ATGAGCTCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 840
Db 869 ATGAGCTCTCTAAGCTGTCACATCTCAAGGGCCACAGAGGGCAAGCCCTCAAGAGC 928
Qy 841 CACAAGTGGATCTCCATCAGAGGGCCCTGCTTCTTCCAGTGTGCTGCAACCCCATC 900
Db 929 CACAAGTGGATCTCCATCAGAGGGCCCTGCTTCTTCCAGTGTGCTGCAACCCCATC 988
Qy 901 CTCTATGCTTCTGAGGCAAACTTAAAGCTCTGCGCAGCAGCTCTGCGCTGCTGCTGCTGCTG 960
Db 989 CTCTATGCTTCTGAGGCAAACTTAAAGCTCTGCGCAGCAGCTCTGCGCTGCTGCTGCTGCTG 1048
Qy 961 AGCAGAGGCTCCAGCTCTCAAGTCTTCCAAAGGAAAGCGGGGTGCAACCTCTTCCGCTC 1020
Db 1049 AGCAGAGGCTCCAGCTCTCAAGTCTTCCAAAGGAAAGCGGGGTGCAACCTCTTCCGCTC 1108
Qy 1021 TCCAGAGGTCAGATCTCCATCAGAGGGCCCTGCTTCTTCCAGTGTGCTGCAACCCCATC 1080
Db 1109 TCCAGTGAAGTCTGAGTCTTCAAGTCTTCACTCCAGCTAA-----CACAGATGTAA 1160
Qy 1081 AAT 1140
Db 1161 GACTTTTTTTTATACGATAAATACCTTTTTTTTAAAGTTACACATTTTCCAGATATAAG 1220
Qy 1141 ACTGACAGCTTGTGACGTTTTTTTTTTTTTTTAAATGACTGTGTTGGGAGTTTATGTTCC 1200
Db 1221 ACTGACCAATATGTTACAGTTTTTTTATGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1274
Qy 1201 TCTAGTTTTTGTGAGGTTGACTTATATATATATATATATATATATATATATATATATATAT 1260
Db 1275 TTTAGTTTTTGTGAGGTTTAAATGACTTATATA-----TAAATTTTTTTTGTTCATAT 1330
Qy 1261 GAATGAGCGTCTAGGAGGACCTGTGCGCAAGTTCTTAGTAGCTGTGTTATCTGTGTAG 1320
Db 1331 TGATGTGTGCTAGGAGGACCTGTGCGCAAGTTCTTAGTTGCTGTATCTCTCGTGTAG 1390
Qy 1321 GACTGTAGAACTGTAGAGGAGAACTGACATCTCAGATATGTTGGTAAATTAATAA 1380
Db 1391 GACTGTAGAA-----AAGGGAATGAACATTCAGAGCGGTGTAGTGAATCACTGATAA 1442

RESULT 5

US-10-342-887-912
; Sequence 912, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9101-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIORITY FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIORITY FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIORITY FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIORITY FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-912

Query Match 55.1%; Score 933.4; DB 13; Length 1679;
Best Local Similarity 81.8%; Pred. No. 2.1e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

Qy 1 ATATACATCTCTGATACACTCTCTGAAGAGTGGGTCTGGAGACTATGACTCCAAACAAG 60
Db 104 ATATACATCTCTGATACACTCTCAAGAGAAATGGCTCAGGGGACTATGACTCCATGAAG 163
Qy 61 GAACCCCTGTTCCCGGATGAAACGTCATTTCAATAGGATTTCTTCCGCCACCATCTAC 120
Db 164 GAACCCCTGTTCCCGTGAAGAAATGCTAAATTTCAATAAATCTTCTGCCCAACCATCTAC 223
Qy 121 TTCTCATCTTCTTGAAGTGTCCATATCATCTGCAACTGTCAACCTCTACAGCAGGTTCTC 180
Db 224 TCCATCATCTTCTTAACTGGCAATTTGGGCAATGGATTGGTTCATCTCTGGTTCATGGTTAC 283
Qy 181 CAGAAGAGCTTAAGAGCATGACGCAAGTACCGGCTGCACTGTCTGAGTGGCTGACCTC 240
Db 284 CAGAAGAACTGAGAAGCATGACGCAAGTACAGGCTGCACTGTCTGAGTGGCGACCTC 343
Qy 241 CTCTTTGTATCATCAGTCCCTTCTGGGCAATTTGATGCCATGGCTGATGGTACTTTGGG 300
Db 344 CTCTTTGTATCATCAGCTTCCCTTCTGGGCAATTTGATGCCGCAAACTGTTACTTTGGG 403
Qy 301 AAATTTTGTGTAAGGCTGTCCATATCATCTGCAACTGTCAACCTCTACAGCAGGTTCTC 360
Db 404 AACTTCTATGCAAGGAGTCCATGTCTATACAGTCAACCTCTACAGCAGTGTCTC 463
Qy 361 ATCTGCGCTTCTATCAGCTGAGCCGCTGCTGCGCATTTGTCCAGCCACCAACAGTCAA 420
Db 464 ATCTGCGCTTCTATCAGTCTGAGCCGCTGCTGCGCATTTGTCCAGCCACCAACAGTCAA 523
Qy 421 AGGCCAAGAACTGCTGCTGAAAGGCAAGTCTATGTTGGCGCTCTGATCCCGCCCTC 480
Db 524 AGGCCAAGAACTGTTGGCTGAAAGGCTCTATGTTGGCGCTCTGATCCCGCCCTC 583
Qy 481 CTCTGACTATCTGACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

584 CTGCTGACTATTCGCGACTTCATCTTGGCAAG-----TCAGTGAGGCA 628
541 GATCAGAGGTACATCTGTGCGCGCTTTACCCGATAGCCTGTGGATGGTGTTCAA 600
629 GATGACAGATATCTGTGACCGCTTCTACCCCAATGACTTGGGTGGTGTGTTCCAG 688
601 TTCCAGCATATAATGGTGGGTCTCATCTGCGCGGCAATGCTATCTCTCTTTACTGC 660
689 TTTGAGCAGCATCATGGTGGGCTTATCTGCTGGTATTTGCTATCTCTGCTCTATTGC 748
661 ATCATCATCTTACAGTGTACATCTCAAGGCGCCAGAGCGCGAGGCGCTTCAAGAGC 720
749 ATTATCATCTTCAAGTGTACATCTCAAGGCGCCAGAGCGCGAGGCGCTTCAAGAGC 808
721 ACAGTCACTCTCATCTAGCTTTCTTTGCGCTGTGCTGCTCATATATATGCGGATCAGC 780
809 ACAGTCACTCTCATCTAGCTTTCTTGGCTGTGCTGCTTACTACATGCGGATCAGC 868
781 ATCAGCTCTCTCATCTTTGGGAGTATCATCAAGGAGTGTGACTTTCGAGAGCATTTG 840
869 ATCAGCTCTCTCATCTTTGGGAGTATCATCAAGGAGTGTGACTTTCGAGAGCATTTG 928
841 CACAAGTGGATCTCCATCACAGAGGCGCTTGGCTTTCTTCCACTGTTGCTTGAACCCCATC 900
929 CACAAGTGGATCTCCATCACAGAGGCGCTTGGCTTTCTTCCACTGTTGCTTGAACCCCATC 988
901 CTATATGCTCTTCTGCGGCGCAAGTCTCAAGAGTCTTCCAGCATGCGACTCACTCATC 960
989 CTATATGCTCTTCTTGGAGCAAAATTTAAACCTCTGCGGAGCGACTCACTCTGTG 1048
961 AGCAGAGGCTCAGCTCAAGATCTTTCCAAAGGAAAGGGGTGACACTCTTCTGCTC 1020
1049 AGCAGAGGCTCAGCTCAAGATCTTTCCAAAGGAAAGGGGTGACACTCTTCTGCTC 1108
1021 TCCAGGAGTCAAGATCTTCCAGTCTTCACTCCAGCTTAACCCCTTATGCAAGAGTATAT 1080
1109 TCCAGTCACTGAGTCTTCAAGTCTTCACTCCAGCTTAACCCCTTATGCAAGAGTATAT 1160
1081 AATATATATATATATGATTAAGAACTTTTATGTTTACACATTTTCCAGATATAAGAG 1140
1161 GACTTTTTTTATGATTAAGTAACTTTTATGTTTACACATTTTCCAGATATAAGAG 1220
1141 ACTGACAGTCTGTACAGTTTTTTTTTTTTTTTTTAAATGAGTGTGGAGTTTATGTTCC 1200
1221 ACTGACCAATATGTACAGTTTTTATGCTTGTGGATTTTGT-----CTTGTGTTTC 1274
1201 TCTAGTTTTTGTGAGTTTGAATTTATATAATAATTTGTTTGTGTTTGTTCATGT 1260
1275 TTTAGTTTTTGTGAGTTTGAATTTATGACTTATTTATA-----TAAATTTTTTGTGTTTCATAT 1330
1261 GAATGAGGCTGTAGGAGGAGCTGTGGCCAGTCTTAGTAGCTGTTTATCTGTGTAG 1320
1331 TGATGTGTGTCTAGGAGGAGCTGTGGCCAGTCTTAGTAGCTGTTTATCTGTGTAG 1390
1321 GACTGTAGAACTGTAGAGGAGAACTGAACTTCCAGATGTGTGTAATTAAGATAAA 1380
1391 GACTGTAGAA-----AAGGGAAGTGAATTTCCAGAGGCTGTAGTAATCACTGATAA 1442
1381 GCTAGCGGTGATCTCTCAGCTGTGTGTCATA 1411
1443 GCTAGAAATGATCTCCAGCTGTTTATGATA 1473

RESULT 6

US-10-151-274-14
; Sequence 14, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeek, Tne

; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-151-274-14

Query Match 55.1%; Score 933.4; DB 13; Length 1679;

Best Local Similarity 81.8%; Pred. No. 2.1e-236;

Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACATCTCTGATAACTACTCTCAAGAAGTGGGGTCTGGAGACTATGACTTCCAAACAG 60
DB 104 ATATACATCTCTGATAACTACTCAAGAGGAATGGGCTCAGGGAGTATGACTTCCATGAG 163
QY 61 GAACCCCTGCTTCCGGGATGAAAACTCCATTTCAATAGGATCTTCTCCGCCACCATCTAC 120
DB 164 GAACCCCTGCTTCCGGGATGAAAACTCCATTTCAATAGGATCTTCTCCGCCACCATCTAC 223
QY 121 TTATCATCTCTTTCGACTGCGCATAGTCGCAATGGATTGGTGTGATCCTGCTCATGGTTAC 180
DB 224 TCCATCATCTCTTTCGACTGCGCATAGTCGCAATGGATTGGTGTGATCCTGCTCATGGTTAC 283
QY 181 CAGAGAGAGCTTAAGGAGCATGACGAGCAAGTACCGGCTGACACCTGTCACTGAGTGGCTGACCTC 240
DB 284 CAGAGAGAGCTTAAGGAGCATGACGAGCAAGTACCGGCTGACACCTGTCACTGAGTGGCTGACCTC 343
QY 241 CTCTTTGCTATCACACTCCCTTCTGGGCAAGTGTGATGCGATGGCTGAGTGTGACTTTGGG 300
DB 344 CTCTTTGCTATCAGCCTTCCCTTCTGGGCAAGTGTGATGCGGCTGACACCTGTGAGTGTGACTTTGGG 403
QY 301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGAGAGCTTCTC 360
DB 404 AACTTCTTATGCAAGGAGTGTCCATGTCTATACAGTCAACCTCTACAGAGTGTGCTC 463
QY 361 ATCTTGGGCTTATCAGCTGAGACCGGTACCTTGGCCATTTGCGGCAAGTGTGAGTGTGAGTCAA 420
DB 464 ATCTTGGGCTTATCAGCTGAGACCGGTACCTTGGCCATTTGCGGCAAGTGTGAGTGTGAGTCAA 523
QY 421 AGGCCAAGGAACTCTGCTGGAAGGAGTGTATGTTGGGCTGTGATGCGGCTGAGTCCAGCCCTC 480
DB 524 AGGCCAAGGAACTCTGCTGGAAGGAGTGTATGTTGGGCTGTGATGCGGCTGTGATGCGGCTC 583
QY 481 CTCTGACTATATACCTGACTTCACTCTCTGCGGCAAGTGTGAGTGTGAGTGTGAGTCAA 540
DB 584 CTGCTGACTATTTCCCGACTTCACTTTGCCAAG-----TCAGTGAGGCA 628
QY 541 GATGACAGTATCATCTGTGACCGGCTTTACCCCGATGAGCTGTGATGAGTGTGAGTGTGAGTCAA 600
DB 629 GATGACAGTATATCTGTGACCGGCTTTACCCCGATGAGCTGTGATGAGTGTGAGTGTGAGTCAA 688
QY 601 TTCCAGCATATAATGGTGGGTCTCATCTCTGCGGCAAGTGTGAGTGTGAGTGTGAGTCAA 660
DB 689 TTTGAGCAGCATCATGGTGGGCTTATCTGCTGGTATTTGCTATCTCTGCTGCTATTGC 748
QY 661 ATCATCATCTTAAAGTGTGACACTTCCAGGCGCCAGAGCGCGAGGCGCTTCAAGAGC 720
DB 749 ATTATCATCTTCAAGTGTGACACTTCCAGGCGCCAGAGCGCGAGGCGCTTCAAGAGC 808
QY 721 ACAGTCACTCTCATCTAGCTTTCTTTGCGCTGTGCTGCTCATATATATGCGGATCAGC 780
DB 809 ACAGTCACTCTCATCTAGCTTTCTTTGCGCTGTGCTGCTTACTACTATGCGGATCAGC 868
QY 781 ATGAGTCTCTCATCTTTTGGGAGTGTCAAGAGGAGTGTGACTTTCGAGAGCATTTG 840

869 ATGACCTCTTCACTCTCTGGAATCATCAGCAAGGGTGTGAGTTTGAGAACACTGTG 928
841 CACAAGTGGATCTCCATCACAAGAGCCCTCGCTTCTTCCACTGTTCCTGAAACCCGATC 900
929 CACAAGTGGATTTCCATCAGCAGAGCCCTAGCTTCTTCCACTGTGTGTGAAACCCGATC 988
901 CTCTATGCTTCTCTCGGGGGCAAGTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATG 960
989 CTCTATGCTTCTCTGAGGCAATTTAAACCTCTGCCAGCAGCAGCTCACTCTGTG 1048
961 AGCAGAGCTCCAGCTCAGATCCTTTCCAAAGAAAGCGGGTGGACACTCTTCCGTC 1020
1049 AGCAGAGGCTCCAGCTCAGATCCTCTCCAAAGAAAGCGGGTGGACACTCTCTGT 1108
1021 TCCAGGAGTCAGATCCTCCAGCTTTTCACTCCAGCTAAACCTTATGCCAAAGACTTAT 1080
1109 TCCAGTGTGAGTCTTCAAGTCTTCAAGTCTTCACTCCAGCTAA-----CAGAGTGTAAA 1160
1081 AATATATATATATATATATAAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG 1140
1161 GACTTTTTTTTATACGATAAATAACTTTTTTTAAAGTTACACATTTTTCAGATATAAAG 1220
1141 ACTGACAGCTCTGACAGTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTTCC 1200
1221 ACTGACCAATATGTACAGTTTTTATGCTTGTGTGGAATTTTGT-----CTTGTGTTTC 1274
1201 TCTAGTTTTTGTGAGTTTGTACTTAATTTATATATAATTTGTTTTTGTGTTTCATGT 1260
1275 TTTAGTTTTTGTGAGTTTAAATGACTTATTTATA-----TAAATTTTTTGTGTTTCAT 1330
1261 GAATGAGGCTTAGGAGGAGCTGTGGCAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG 1320
1331 TGATGTGTGTCTAGGAGGAGCTGTGGCAAGTTCTTAGTTGTGTGTGTCTCGTGTGTAG 1390
1321 GACTGTAGAACTGTAGAGAGAACTGAACTTCCAGAAATGTGTGTAATTTGAATAAA 1380
1391 GACTGTAGAA-----AAGGAACTGAACTTCCAGAGCGTGTAGTGAATCACGTAAA 1442
1381 GCTAGCGGTGATCTCTAGCTGTTCTGTGCATA 1411
1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

RESULT 7

US-10-172-118-912
; Sequence 912, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172.118
; PRIORITY FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380, 770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 912
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 003467
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-912

Query Match 55.1%; Score 933.4; DB 13; Length 1679;
Best Local Similarity 81.8%; Pred. No. 2.1e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;
QY 1 ATATACACTTCTGATAACTACTCTGAAAGAGTGGGTCTGGAGACTATGACTCCCAACAAG 60
DB 104 ATATACACTTCTGATAACTACTACACAGGAAATGGGCTCAGGGGACTATGACTCCCAACAAG 163
QY 61 GAAACCTCTGCTTCGGGATGAAAACGCTCCATTTCAATAGGATCTTCTCCCAACCACTTAC 120
DB 164 GAAACCTCTGCTTCGGGATGAAAACGCTCCATTTCAATAGGATCTTCTCCCAACCACTTAC 223
QY 121 TTCATCATCTTCTTACCTGGCATAGTCGGCAATGATTTGGTATCTCTGGTATGTTGGTAT 180
DB 224 TCCATCATCTTCTTAACTTGGCATAGTCGGGCAATGATTTGGTATCTCTGGTATGTTGGTAT 283
QY 181 CAGAAGAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTGTCACTGTGCTGCTGACCTC 240
DB 284 CAGAAGAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTGTCACTGTGCTGCTGACCTC 343
QY 241 CTCTTTGTATCACACTCCCTTCTGGGAGTGTATGCCATGCTGGTGTGACTGTGTTGGG 300
DB 344 CTCTTTGTATCACACTCCCTTCTGGGAGTGTATGCCATGCTGGTGTGACTGTGTTGGG 403
QY 301 AAATTTTGTGTAAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGAGAGCTTCTC 360
DB 404 AAATTTTGTGTAAAGGCTGTCCATATCATCTACACTGTCAACCTCTACAGAGAGCTTCTC 463
QY 361 ATCTGGGCTTTCATCAGCTGACGCTGACCGGTACCTGCGCATTTGTCACGCCCAACCAAGTCAA 420
DB 464 ATCTGGGCTTTCATCAGCTGACGCTGACCGGTACCTGCGCATTTGTCACGCCCAACCAAGTCAA 523
QY 421 AGGCCAAGAACTGCTGGCTGAAAGGAGTCTATGTTGGGCTGTGGTGTGCTGCTGCTGCTC 480
DB 524 AGGCCAAGAACTGCTGGCTGAAAGGAGTCTATGTTGGGCTGTGGTGTGCTGCTGCTGCTC 583
QY 481 CTCTGTGACTATACCTGACTTTCATCTTTGCGGAGCTCAGCCAGGGGAGCATCATAGTCAGGGG 540
DB 584 CTCTGTGACTATACCTGACTTTCATCTTTGCGGAGCTCAGCCAGGGGAGCATCATAGTCAGGGG 628
QY 541 GATGACAGTACATCTGTGACCGCTTTACCCCGATAGCTGTGATGTTGGTGTGTTTCAA 600
DB 629 GATGACAGTATATCTGTGACCGCTTTACCCCGATAGCTGTGATGTTGGTGTGTTTCAA 688
QY 601 TTCCAGCATATAATGTTGGGTCTCATCTGCGCGGATCGTCTCATCTCTCTCTGTTACTGTC 660
DB 689 TTTGACCATCATGTTGGCTTATCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 748
QY 661 ATCATCATCTTAAGCTGTCACTCCAAAGGCGCACAGAGCGCAAGGCCCTCAAGAG 720
DB 749 ATTATCATCTCAAGCTGTCACTCCAAAGGCGCACAGAGCGCAAGGCCCTCAAGAG 808
QY 721 ACAGTCATCTCATCTAGCTTCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 809 ACAGTCATCTCATCTAGCTTCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868
QY 781 ATCGACTCTCTCATCTTGGGAGTCACTCAAGAAAGATGTGATTTGAGAGAGCTGTTG 840
DB 869 ATCGACTCTCTCATCTTGGGAGTCACTCAAGAAAGATGTGATTTGAGAGAGCTGTTG 928
QY 841 CACAAGTGTATCTCATCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 929 CACAAGTGTATTTCCATCAGGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 988
QY 901 CTCTATGCTTCTCTCGGGGCAAGTTCAAAAGCTCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 989 CTCTATGCTTCTCTGAGGCGCAATTTAAACCTCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG 1048
QY 961 AGCAGAGGCTCCAGCTCAGATCTCTTCCAAAGAAAGCGGGTGGACACTCTTCCGCTC 1020
DB 1049 AGCAGAGGCTCCAGCTCAGATCTCTTCCAAAGAAAGCGGGTGGACACTCTTCCGCTC 1108
QY 1021 TCCAGGAGTCAGAACTCTCTCCAGTCTTTCACCTCAAGCTTATGCTCAAGAGACTTATAT 1080

Db	1109	TCACATGAGCTCGAGTCTTCAAGTTTTCACCTCAGCTAA-----CACAGATGTAAAA	1160
Qy	1081	AAATATATATATATGATAAAGAACTTTTATGTTACACATTTTCCAGATATAAGAG	1140
Db	1161	GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTACACATTTTCCAGATATAAAG	1220
Qy	1141	ACTGACCAGCTTTGTACAGTTTTTTTTTTTTTAAATGACTGTGGGAGTTTATGTTCC	1200
Db	1221	ACTGACCAATTTGTACAGTTTTTTTATGCTGTGTGGAATTTTGT-----CTTGTGTTTC	1274
Qy	1201	TC'TAGTTTTTGTGAGGTTTGACCTTAATTTATATAAATATTGTTTTTTTGTGTTTTCATGT	1260
Db	1275	TTT'AGTTTTTGTGAAGTTTAATTGACTTATTATA-----TAAATTTTTTTTGTTCATAT	1330
Qy	1261	GAATGACGCTTAGGCGAGGACCTGTGGCCAAAGTTCTTAGTAGCTGTTTATCTGTGTGTAG	1320
Db	1331	TGATGTGTGTCTAGGCGAGGACCTGTGGCCAAAGTTCTTAGTGTGTATGTTCTCGTGGTAG	1390
Qy	1321	GACTGTAGAACTCTAGAGGAAGAAACTGAACATTTCCAGAAATGTGTGTAATTTGAATAAA	1380
Db	1391	GACTGTAGAA-----AAGGNACTGAACATTTCCAGGCGTGTAGTGAATCACGTAAA	1442
Qy	1381	GCTAGCGGTGAATCCTCAGCTGTTGCTGCATA	1411
Db	1443	GCTAGAAATGATCCCGAGCTGTTTATGCATA	1473

RESULT 8

```

US-10-170-385-332
; Sequence 332, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-332

```

	Query Match	55.1%	Score 933.4	DB 13	Length 1679
	Best Local Similarity	81.8%	Pred. No. 2.1e-236		
	Matches 1154	Conservative 0	Mismatches 216	Indels 41	Gaps 5
Qy	1	ATATACACTTCTGTATAACTACTCTGGAAGAAGTGGGGTCTGGAGACTATGACTCCACAAG	60		
Db	104	ATATACACTTCAGATAACTACCCGAGGAATGSGCTCAGGGACTATGACTCCATGAAG	163		
Qy	61	GAACCCGTCTCCGGATGAAAAGTCGATTTCAATAGGATCTTCCTGCCACCACATCTAC	120		
Db	164	GAACCCGTCTTCGGTGAAGAAAATGCTAAATTCATAAAATCTTCCTGCCACCACATCTAC	223		
Qy	121	TTTCATCATCTTCTTTGACTGGCATAGTCGGCAATGGATTTGGTGTGATCCTCGGTTCATGGGTTC	180		

APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1el Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-021-660-58

Query March 55.1%; Score 933.4; DB 15; Length 1679;
Best Local Similarity 81.8%; Pred.No.2.e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGATACTCTGAAAGAGTGGGGCTGGAGACTATGACTCCAAACAG 60
DB 104 ATATACACTTCTGATACTCTGAAAGAGTGGGGCTGGAGACTATGACTCCAAACAG 163
QY 61 GAACCTGCTTCCGGGAGTAAAGCTCCATTTCAATAGGATCTTCCGCCCAACATCTAC 120
DB 164 GAACCTGCTTCCGGGAGTAAAGCTCCATTTCAATAGGATCTTCCGCCCAACATCTAC 223
QY 121 TCCATCATCTTCTGACTGGCATAGTCGGCAATGGATTTGGTGTCTCTGGTGCATGGGTTAC 180
DB 224 TCCATCATCTTCTGACTGGCATAGTCGGCAATGGATTTGGTGTCTCTGGTGCATGGGTTAC 283
QY 181 CAGAAGAGCTAAGAGCATGACGCAAGTACCGGTGACCTGTGACGTGGGTGACCTC 240
DB 284 CAGAAGAGCTAAGAGCATGACGCAAGTACCGGTGACCTGTGACGTGGGTGACCTC 343
QY 241 CTCCTTGTCTACACACTCCCTCTGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
DB 344 CTCCTTGTCTACACACTCCCTCTGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 403
QY 301 AAATTTTGTGTAAAGGTGTCCATATCATCTACACTGTCAACCTCTACACAGCGTTCTC 360
DB 404 AACTTCTATGCAAGGAGTCCATGTCTATCTACACAGTCAACCTCTACACAGTGTCTC 463
QY 361 ATCTGGCTTTCATCAGCTGGACGGTACTCTGCCATTTGTCCACGCCCAACAGTCAA 420
DB 464 ATCTGGCTTTCATCAGCTGGACGGTACTCTGCCATTTGTCCACGCCCAACAGTCAA 523
QY 421 AGGCAAGGAACTGCTGGCTGAAAGGCACTCTATGTGGCGCTCTGGATCCCGACCTC 480
DB 524 AGGCAAGGAACTGCTGGCTGAAAGGCACTCTATGTGGCGCTCTGGATCCCGACCTC 583
QY 481 CTCCTGACTATACCTGACTTTCATCTTTGGCGACGTGACGCCAGGGGGACATCATCTCAGGGG 540
DB 584 CTGCTGACTATTCCTGACTTTCATCTTTGGCGACGTGACGCCAGGGGGACATCATCTCAGGGG 628
QY 541 GATCAGAGTATCATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTGGTGGTCAA 600
DB 629 GATCAGAGTATCATCTGTGACCGCTTTACCCGATAGCTGTGGATGGTGGTGGTGGTCAA 688
QY 601 TTCAGCATATTAATGGTGGTTCATCTCTGCCGGCATCTGATCTCTCTCTCTCTCTCTCTCT 660
DB 689 TTTTCAGCATATTAATGGTGGTTCATCTCTGCCGGCATCTGATCTCTCTCTCTCTCTCTCTCT 748
QY 661 ATCATCATCTTAAGCTGTACACTCCCAAGGCCCAAGGCGCAGAGCGCCCTCAAGAGC 720
DB 749 ATTATCATCTCAAGCTGTACACTCCCAAGGCCCAAGGCGCAGAGCGCCCTCAAGAGC 808
QY 721 ACAGTCTATCTCTATCT 780

DB 809 ACAGTCATCTCTCATCTCTGGCTTTCTTCGCCCTGTGGCTGCCTTACTACATTTGGGATCAGC 868
QY 781 ATCGAGCTCTCTCATCTCTTTGGGAGTCTCATCAAGCAAGGATGTGACTTCGAGAGCAATTTGTG 840
DB 869 ATCGAGCTCTCTCATCTCTCTCTGAAATCATCAAGCAAGGATGTGACTTCGAGAGCAATTTGTG 928
QY 841 CACAAGTGGATCTCCATCAGAGAGCCCTCGCTTTCTTCCACTGTGTGCTGCTGAAACCCCATC 900
DB 929 CACAAGTGGATTTCCATCAGAGAGCCCTAGCTTTCTTCCACTGTGTGCTGCTGAAACCCCATC 988
QY 901 CTCATGCTCTCTCGGGGCAAGTTCAAAGCTCTGCCAGCATGCTGCCAGCATGCACTCAACTCCATG 960
DB 989 CTCATGCTCTCTCGGGGCAAGTTCAAAGCTCTGCCAGCATGCTGCCAGCATGCACTCAACTCCATG 1048
QY 961 AGCAGAGCTCCAGCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTTCCTGTC 1020
DB 1049 AGCAGAGCTCCAGCTCAAGATCCTTTCCAAAGGAAAGCGGGGTGGACACTCTTTCCTGTC 1108
QY 1021 TCCAGGAGTCCAGATCCTCCAGTTTTCACCTCCAGCTAACCTTATGCAAGAGACTTATAT 1080
DB 1109 TCCAGTGTGTGAGTCTCAAGTTTTCACCTCCAGCTAA-----CACAGATGTAAAA 1160
QY 1081 AATATATATATATATATGATAAGAACTTTTATTTATGTTACACATTTTCCAGATATAAGAG 1140
DB 1161 GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTTACACATTTTTCAGATATAAAG 1220
QY 1141 ACTGACGAGTCTGTACAGTTTTCACCTTTTAAATGACTGTGGAGTTTATGTTCC 1200
DB 1221 ACTGACCAATATGTACAGTTTTCACCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 1274
QY 1201 TCTAGTCTTTGTGAGCTTTGACTTAAATTTATATATAATTTCTTTTCTTTTCTTTTCTTTTCTTT 1260
DB 1275 TTTAGTCTTTGTGAGTTTAAATGACTTATTTATA---TAAATTTTTTTTCTTTTCTTTTCTTT 1330
QY 1261 GAATGAGGCTGTAGCAGGACCTGTGGCCAAAGTCTTTAGTAGCTGTGTTATCTGTGTGTAG 1320
DB 1331 TGATGTGTGTCTAGCAGGACCTGTGGCCAAAGTTCTTTAGTTGCTGTATGTCTCTCGTGTAG 1390
QY 1321 GACTGTAGAACTGTAGAGGAAAGAACTGAAACATTCAGAAATGTGTGTAATTTGAATAAA 1380
DB 1391 GACTGTAGAA-----AAGGAACTGAAACATTCAGAGCGGTGTAGTGAATCAGCTAAA 1442
QY 1381 GCTAGCCGTGATCTTCAGCTGTGTGTCATA 1411
DB 1443 GCTAGAAATGATCCCGACGTGTTTATGATA 1473

RESULT 11
US-10-341-434-177
; Sequence 177, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 177
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(1144)
; OTHER INFORMATION:
US-10-341-434-177

Query Match 55.1%; Score 933.4; DB 16; Length 1679;
Best Local Similarity 81.8%; Pred. No. 2.1e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGATTAACACTCTGGAAGAGTGGGCTCTGGAGACTATGACTCCCAACAG 60
DB 104 ATATACACTTCTGATTAACACTCTGGAAGAGTGGGCTCTGGAGACTATGACTCCCAACAG 163
QY 61 GAACCCCTGCTTCGGGATGAAGAAAGTCCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 120
DB 164 GAACCCCTGCTTCGGGATGAAGAAAGTCCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 223
QY 121 TTCAATCATCTTCTGACTGTCATAGTTCGGCAATGGATGGTGTATCTGCTCATGGGTAC 180
DB 224 TCCATCATCTTCTTAAGTGGCAATGGTGTATCTGCTCATGGGTAC 283
QY 181 CAGAAGAGTGAAGAGCATGACGCAAGTACCGGCTGCACCTGTCACTGGGTGACCTC 240
DB 284 CAGNAGAACTGAGAGCATGACGCAAGTACAGGCTGCACCTGTCACTGGGTGACCTC 343
QY 241 CTCTTTGTCATACATCTCCCTTCTGGGCAAGTGTATGCGCATGGCTGACTGTGTTGGG 300
DB 344 CTCTTTGTCATACATCTCCCTTCTGGGCAAGTGTATGCGCATGGCTGACTGTGTTGGG 403
QY 301 AATTTTGTGTAAGGTGTCATATCATCTACATCTCAACCTCTACAGCAGGTTCTC 360
DB 404 AACTTCTATGCAAGGAGTCCATGTCATCTACAGTCACTCTACAGCAGTGTCTC 463
QY 361 ATCTGCGCTTCTATGAGCTGGAACGCTACTCTGCGCATTTCCACGCCCAACAGTCAA 420
DB 464 ATCTGCGCTTCTATGAGCTGGAACGCTACTCTGCGCATTTCCACGCCCAACAGTCAA 523
QY 421 AGGCAAGGAACTGCTGGAAGGAGTCTATGTTGGGCTGCTGGATCCCGACGCTC 480
DB 524 AGGCAAGGAACTGCTGGAAGGAGTCTATGTTGGGCTGCTGGATCCCGACGCTC 583
QY 481 CTCCTGACTATACCTGACTTCTATCTTTCGAGCTGACGCCAGGGGACATCACTCAGGGG 540
DB 584 CTGCTGACTATTCGAGCTTCTATCTTTCGCAAG- - - - -TCAGTGAGGCA 628
QY 541 GATGACAGGTATCTGAGCGGCTTACCCGATAGCTGTGGATGGTGGTGTTCAA 600
DB 629 GATGACAGATATCTGAGCGGCTTACCCGATAGCTGTGGGTTGGTGTTCAG 688
QY 601 TTCAGGATATAAAGTGGGCTCATCTGCGCGGCTGCTCATCTCTCTGTTACTGC 660
DB 689 TTTGAGCATATGTTGGGCTTATCTGCTGCTGATGCTATCTCTCTGCTATTC 748
QY 661 ATCATCATCTTAAGCTGTACATCTCAAGGGGCCACCAAGAGCGCAAGGCCCTCAAGAG 720
DB 749 ATTATCATCTCAAGCTGTACATCTCAAGGGGCCACCAAGAGCGCAAGGCCCTCAAGACC 808
QY 721 ACAGTCATCTCATCTAGCTTCTTTCCTGCTGCTGCTGCTATTTATGTTGGGATCAGC 780
DB 809 ACAGTCATCTCATCTAGCTTCTTTCCTGCTGCTGCTGCTATTTATGTTGGGATCAGC 868
QY 781 ATCGACTCTCTTCTTTCGGGAGTCAATCAAGAGGATGCTATCGAGAGCATTTGTG 840
DB 869 ATCGACTCTCTTCTTTCGGGAGTCAATCAAGAGGATGCTATCGAGAGCATTTGTG 928
QY 841 CACAAGTGGATCTCATACAGAGGCTCTGCTTCTTTCAGTCTGCTGCTGCTGCTGCTGCT 900
DB 929 CACAAGTGGATCTCATACAGAGGCTCTGCTTCTTTCAGTCTGCTGCTGCTGCTGCTGCT 988
QY 901 CTCTATGCTTCTCTGCGGCGCAAGTTCAAAAGCTCTGCGGAGCATGCACTCAATCCCATG 960
DB 989 CTCTATGCTTCTTTCGGAGCCAAATTTAAACCTCTGCGGAGCAGCAGTCACTCTGCTG 1048
QY 961 AGCAGAGGCTCAGGCTCAAGATCTTTCGAAAGGAGGAGGAGTGTGACACTCTTCTGCTC 1020
DB 1049 AGCAGAGGCTCAGGCTCAAGATCTTTCGAAAGGAGGAGGAGTGTGACACTCTTCTGCT 1108
QY 1021 TCCACGGAGTCAAGATCTCTCAGATTTTTCACCTCCAGTAAACCTTTATGCAAGACTTATAT 1080

DB 1109 TCCACTGAGTCTGAGTCTTCAAGTTTTTCACTCCAGCTAA- - - - -CACAGATGTAAA 1160
QY 1081 AAT 1140
DB 1161 GACTTTTTTATACGATAAATAACTTTTTTTTAAAGTTACACATTTTTCAGATATAAAG 1220
QY 1141 ACTGACGCTTCTGACAGTTTTTTTTTTTTTTTTTAA-TGACTGTGGAGTTTATGTTCC 1200
DB 1221 ACTGACCAATATTTACAGTTTTTTTATGCTTGTGGATTTTTGT- - - - -CTTGGTTTC 1274
QY 1201 TCTAGTTTTTCTGAGGTTTTGACTTTAAATTTATATATAATTTGTTTTTTTGTGTTTCATGT 1260
DB 1275 TTTAGTTTTTCTGAGTTTTTAATGACATTTATATA- - - - -TAAATTTTTTTTGTTCATAT 1330
QY 1261 GAATGAGGCTTCTAGCAGGACCTGTGGCAAGTCTTTAGTAGCTGTTTATCTGTGTAG 1320
DB 1331 TCAATGTGTCTAGGCAAGCTGTGGCAAGTCTTTAGTTGCTGTATGTCTCTGTGGTAG 1390
QY 1321 GACTGTGAAGCTGTAGAGGAAGAACTGAACATTTCCAGAAATGTGTGGTAAATTTGAATAA 1380
DB 1391 GACTGTGAAG- - - - -AAGGAACTGAACATTTCCAGAGGCTGTAGTGAATCAGGTAA 1442
QY 1381 GCTAGCCGTGATCTCTAGCTCTTCTGCTGATA 1411
DB 1443 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1473

RESULT 12
US-09-971-392-20
; Sequence 20, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: 2001-10-03
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Template ID: 284616.2
; NAME/KEY: unsure
; LOCATION: 1706
; OTHER INFORMATION: a, t, c, g, or other
US-09-971-392-20

Query Match 55.1%; Score 933.4; DB 10; Length 1711;
Best Local Similarity 81.8%; Pred. No. 2.2e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

QY 1 ATATACACTTCTGATAAATACTCTGGAAGAGTGGGCTCTGGAGACTATGACTCCCAACAG 60
DB 139 ATATACACTTCTGATAAATACTCTGGAAGAGTGGGCTCTGGAGACTATGACTCCCAACAG 198
QY 61 GAACCCCTGCTTCGGGATGAAGAAAGTCCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 120
DB 199 GAACCCCTGCTTCGGGATGAAGAAAGTCCATTTCAATAGGATCTTCTGCGCCACCACTCTAC 258
QY 121 TTTCAATCATCTTCTGACTGTCATAGTTCGGCAATGGATGGTGTATCTGCTCATGGGTAC 180
DB 259 TCCATCATCTTCTTAACTGGCAATTTGGGCAATGGATGGTGTATCTGCTCATGGGTAC 318
QY 181 CAGAAGAGTCAAGGAGCATGACGCAAGAGTACCGGCTGCACCTGTCTAGTGGCTGACCTC 240

Db	1366	TGATGTGTGCTAGGCAGGACCTGTGGCCAAAGTTCTTAGTTGCTGTATGTCTGCTGGTAG	1422
Qy	1321	GACTGTAGAACTGTAGAGGAAGAAACTGAACATTCAGAAATGTGTGGTAAATTTGAATAAA	1380
Db	1426	GACTGTAGAA-----AAGGGAAGCTGAACATTCAGAGCGTGTAGTGAATCACTGATAA	1477
Qy	1381	GCTAGCGGTGATCCTCAGCTGTTGCTGCATA	1411
Db	1478	GCTAGAAATGATCCCGACGCTGTTTATGCATA	1508
RESULT 13			
US-10-101-510-459			
; Sequence 459, Application US/10101510			
; Publication No. US20030148295A1			
; GENERAL INFORMATION:			
; APPLICANT: WAN, JACKSON			
; APPLICANT: WANG, YIXIN			
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE			
; FILE REFERENCE: 15117.0012			
; CURRENT APPLICATION NUMBER: US/10/101,510			
; CURRENT FILING DATE: 2002-03-20			
; PRIOR APPLICATION NUMBER: 60/276,947			
; PRIOR FILING DATE: 2001-03-20			
; NUMBER OF SEQ ID NOS: 805			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 459			
; LENGTH: 1711			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: modified_base			
; LOCATION: (1706)			
; OTHER INFORMATION: a, t, c, g, other or unknown			
US-10-101-510-459			
Query Match 55.1%; Score 933.4; DB 15; Length 1711;			
Best Local Similarity 81.8%; Pred. No. 2.2e-236;			
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;			
Qy	1	ATATACATCTCTGATAACTACTCTGGAAGTGGGTCTGGAGACTATGACTCCCAACAG	60
Db	139	ATATACATCTCAGATAACTATACACCGAGGAAATGGGCTCAGGAGACTATGACTCCATGAAG	198
Qy	61	GAACCCCTGTCTCGGGATGAAACGTCCTATTTCAATAGGATCTTCTGCCCCCACCATTCTAC	120
Db	199	GAACCCGTGTTCCGTGAAGAAATGCTAATTTCAATAAAATCTTCTGCCCCACCATTCTAC	258
Qy	121	TTCATCATCTTCTGACTGGCATAGTCGGCANTGGAATTGGTGATCTCTGTCATGGGTTAC	180
Db	259	TCCATCATCTTCTTAACTGGCATTTGTGGCAATGGATTGGTTCATCTCTGTCATGGGTTAC	318
Qy	181	CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTGCACCTGTCACTGGTGGTGACCTC	240
Db	319	CAGAAGAACTGAGAGCATGACGGACAAGTACAGGCTGCACCTGTCACTGGCGGACCTC	378
Qy	241	CTCTTTGTCAACACACTCCCTTCTGGGCAGTGTGATGCATGGCTGACTGGTACTTTGGG	300
Db	379	CTCTTTGTCAACACGCTTCCCTTCTGGGCAGTGTGATGCCTGGCAAACTGGTACTTTGGG	438
Qy	301	AAATTTTCTGTAAGCTCTCCATATCATCTACACTGTCAACCTCTACAGCAGCGTTCTC	360
Db	439	AACTTCTTATGCAAGCAGTCATGTCACTACAGTCAACTCTACAGCAGTGTCTTC	498
Qy	361	ATCCTGGCCCTTCATCAGCCTGGACCGGTACCTCGCCATTGTCCACGCCACCAACAGTCAA	420
Db	499	ATCCTGGCCCTTCATCAGTCTGGACCGCTACCTGGCCATCGTCCAGGCCACCAACAGTCAG	558
Qy	421	AGGCCAAGAAACTGTGCTGAAAGGAGCTATGTGGGGCGTCTGGATCCAGCCCTC	480
Db	559	AGGCCAAGAAAGCTGTGTGCTGAAAGGTTGGTCTATGTTGGCGTCTGGAATCCCTGCCCTC	618
Qy	481	CTCCTGACTATACCTGACATTCATCTTTGCCGACGTCAGCCAGGGGGAATCACTCAGTCAGGG	540

```
Db 619 CTGCTGACTATCCCGACTTCACTTTGCCAAG-----TCACTGAGGCA 663
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCGATAGCCTGTGGATGGTGGTGTTCAA 600
Db 664 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTGTGTCCAG 723
Qy 601 TTCAGACATATAATGTGGGTCTCATCTGCGCGGATCGCTCATCTCTCTCTCTACTGC 660
Db 724 TTTTCAGACATCATGGTGGCCATCTGCTGGTATGTTCTCATCTCTCTCTCTATTC 783
Qy 661 ATCATCATCTTAAGCTGTACATCTCAAGGGCCACCAAGAGCGCAAGCGCCCTCAAGAG 720
Db 784 ATTATCATCTCCAGCTGTACATCTCAAGGGCCACCAAGAGCGCAAGCGCCCTCAAGAG 843
Qy 721 ACAGTCACTCTCACTTCAGTCTTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 844 ACAGTCACTCTCATCTGCTGCTTCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Qy 781 ATCGACTCTTCTCATCTTTTGGGAGTCATCAAGCAAGGATGTGACTTTGAGAGCATTTGT 840
Db 904 ATCGACTCTTCTCATCTCTCTGGAATCATCAAGCAAGGATGTGAGTTTGAGAACACTGT 963
Qy 841 CACAGGTGGATCTCCATCAAGAGGCTCTGCTTCTTCCACTGTTGCTGCAACCCCATC 900
Db 964 CACAGGTGGATTTCCATCAAGAGGCTCTGCTTCTTCCACTGTTGCTGCAACCCCATC 1023
Qy 901 CTCTATGCTTCTCTCGGGGCAAGTTCAAAGCTCTGCCAGCATGCACTCAACTCCATG 960
Db 1024 CTCTATGCTTCTCTGAGCAAAATTAAGCTCTGCCAGCATGCACTCAACTCTGTG 1083
Qy 961 AGGAGGCTCCAGCTCAAGATCTTCCAAAGAAAGCGGGTGGACACTCTTCGCTC 1020
Db 1084 AGGAGGCTCCAGCTCAAGATCTTCCAAAGAAAGCGGGTGGACACTCTTCGCTC 1143
Qy 1021 TCCAGGAGTCAGATCTCTCAGCTTTTCACTCCAGCTTAACCTTATGCAAGACTTAT 1080
Db 1144 TCCAGTGTGCTGAGTCTTCAAGTTTCACTCCAGCTAA-----CACAGATGTAAA 1195
Qy 1081 AATATATATATATATGATAAGAACTTTTATGTTACATATTTTCCAGATATAGAG 1140
Db 1196 GACTTTTTTTTATACGATAAATACTTTTTTTTAAAGTTACACATTTTTCAGATATAA 1255
Qy 1141 ACTGACGAGTCTGACAGTTTTTTTTTTTTTTTTTTTAAATGACTGTGGGATTTATGTC 1200
Db 1256 ACTGACCAATATGTAAGTTTTTATGCTGTGGATTTTGT-----CTTGTGTTTC 1309
Qy 1201 TCTAGTTTTTGTGAGTTTGAATTTATATAAATATTTTGTGTTTTTGTGTTTGTATG 1260
Db 1310 TTTAGTTTTTGTGAAGTTAATGACTTATTTATA-----TAAATTTTTTTTGTTCAT 1365
Qy 1261 GAATGAGGCTTAGGAGGACCTGTGGCAAGTTCTTAGTAGCTGTTTATCTGTGTAG 1320
Db 1366 TGATGTGTGCTAGGAGGACCTGTGGCAAGTTCTTAGTTGTGTGTGTCTGTGTAG 1425
Qy 1321 GACTGTAGAACTGTAGAGGAGAACTGAACATTCAGAAATGTGTGTAATTTGAATAA 1380
Db 1426 GACTGTAGAA-----AGGGAAGTGAACATTCAGAGCGGTAGTGAATCACGTAA 1477
Qy 1381 GCTAGCGGTGATCTCAGCTGTTGCTGCATA 1411
Db 1478 GCTAGAAATGATCCCGAGCTGTTTATGCATA 1508
```

RESULT 14

US-09-104-063-3

; Sequence 3, Application US/09104063

; Patent No. US20020168356A1

; GENERAL INFORMATION:

; APPLICANT: Lee, James

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: PF4A Receptors

; NUMBER OF SEQUENCES: 6

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,063
; FILING DATE: 24-June-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/664228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1737 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-104-063-3
```

Query Match 55.1%; Score 933.4; DB 9; Length 1737;

Best Local Similarity 81.8%; Pred. No. 2.2e-236;

Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy 1 ATATACACTTCTGATAACTACTCTGAAGAAGTGGGCTCTGGAGACTATGACTCCAAAG 60
Db 106 ATATACACTTCTGATAACTACTCACCGAGGAATGGCTCAGGGGACTATGACTCCATGAAG 165
Qy 61 GAACCTGCTCCGGGATGAACCGTCCATTTCATAGGATCTTCCTGCCACCATCTAC 120
Db 166 GAACCTGCTTCCTCGTGAAGAAATGCTAAATTCATAAAATCTTCCTGCCACCATCTAC 225
Qy 121 TTCATCATCTTCTTGACTGGCATAGTCGGCAATGGATTGGTGATCCTGGTCTATGGGTTAC 180
Db 226 TCCATCATCTTCTTAACCTGGCATTTGTGGCAATGGATTGGTCTATGGGTTAC 285
Qy 181 CAGAAAGAGCTAAGGAGCATGACGACAAGTACCGGCTGCAACCTGTCTAGTGGCTGACCTC 240
Db 286 CAGAAAGAACTGAGAAGCATGACGACAAGTACAGGCTGCAACCTGTCTAGTGGCGACCTC 345
Qy 241 CTCCTTGTTCATCACACTCCCTTCCTGGGCACTGATGCGGCTGACTGGTACTTTGGG 300
Db 346 CTCCTTGTTCATCACCGCTTCCCTTCCTGGCACTGATGCGGCTGACTGGTACTTTGGG 405
Qy 301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAAACCTCTACAGAGGTTCTC 360
Db 406 AACTTCCTATGCAAGGCACTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTC 465
Qy 361 ATCTGGGCTTCATCAGCTGACCGGTACCTGCGCCATCTGTCACCGCCACCAAGTCAA 420
```

```
Db 466 ATCTGGCCCTTCATCAGTCTGGACCGCTACTCTGCCATCTGTCCACGCCACCAAGTCAG 525
Qy 421 AGGCCAAGAACTGCTGGCTGAAAGAGCACTATCTGTGGCGCTGTGGATCCAGCCCTC 480
Db 526 AGGCCAAGAACTGCTGGCTGAAAGAGCACTATCTGTGGCGCTGTGGATCCAGCCCTC 585
Qy 481 CTCCTGACTATACCTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 586 CTGCTGACTATCCCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 630
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTC 600
Db 631 GATGACAGATATATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTC 690
Qy 601 TTCAGCATATAGTGGTGGTCTCATCTCTGCGCGCATCTCTCTCTCTCTCTCTCTCT 660
Db 691 TTTGAGCACAATCATGTTGGCTTATCTGCTGCTTATCTGCTCTCTCTCTCTCTCT 750
Qy 661 ATCATCATCTCTAAGCTGTCTACATCTCAAGGGCCACCAAGGGCCCAAGGGCCCTCAAG 720
Db 751 ATTATCATCTCAAGCTGTCTACATCTCAAGGGCCACCAAGGGCCCAAGGGCCCTCAAG 810
Qy 721 ACAGTCATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 811 ACAGTCATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 870
Qy 781 ATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Db 871 ATCGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930
Qy 841 CACAAGTGGATCTCCATCAGAGGGCCCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 931 CACAAGTGGATCTCCATCAGAGGGCCCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 990
Qy 901 CTCTATCGCTCTCTCGGGCCAGTTCAGAAAGCTCTGCCAGCATCACTCACTCCATG 960
Db 991 CTCTATCGCTCTCTCGGGCCAGTTCAGAAAGCTCTGCCAGCATCACTCACTCCATG 1050
Qy 961 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAGCGGGGTGACACTCTCTCCGTC 1020
Db 1051 AGCAGAGGCTCCAGCTCAAGATCTCTTCCAAAGGAAGCGGGGTGACACTCTCTCGIT 1110
Qy 1021 TCCAGGAGTCCAGATCTCTCGAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db 1111 TCCAGTCTGAGTCTCTCAAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1162
Qy 1081 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1140
Db 1163 GACTTTTTTTTATACGATAAATAACTTTTTTTTAAAGTTACACTTTTTCAGATATAAG 1222
Qy 1141 ACTGACAGTCTCTGACAGTTTTTTTTTTTTTTTTTTTTTAAATGACTGTGGAGTTTATGTC 1200
Db 1223 ACTGACCAATATGTAAGTTTTTTTATGCTGTGGATTTTTTGT-----CTTGTGTTTC 1276
Qy 1201 TCTAGTTTTTGTGAGGTTTCTGACTTAATTTATATAATATATTTTTTTTGTGTTTCTCATGT 1260
Db 1277 TTTAGTTTTTGTGAGTTTATGACTTATTTATA-----TAAATTTTTTTTGTGTTTCTAT 1332
Qy 1261 GAATGAGCTCTAGGAGGACCTGTGGCCAAAGTTCTTAGTAGTGTGTTTATCTGTGRTAG 1320
Db 1333 TGATGTGTCTAGGAGGACCTGTGGCCAAAGTTCTTAGTTGCTGTATGCTCTGTGGTAG 1392
Qy 1321 GACTGTAGAACTGTAGAGGAGAACTGTAACATTCAGATATGCTGTAATTAATTAATA 1380
Db 1393 GACTGTAGAA-----AAGGAACTGAAACATTCAGAGCGGTGTAGTAATCACGTAAA 1444
Qy 1381 GCTAGCCGTGATCTCAGCTGTGCTGCATA 1411
Db 1445 GCTAGAAATGATCCCAGCTGTTTATGCATA 1475
```

RESULT 15

```
US-10-666-689-3
; Sequence 3, Application US/10666689
; Publication NO. US20040037830A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Human P4A Receptors, Nucleic Acid Encoding and
; TITLE OF INVENTION: Antibodies Binding Thereto
; FILE REFERENCE: P07062C2D2C1
; CURRENT APPLICATION NUMBER: US/10/666,689
; CURRENT FILING DATE: 2003-09-19
; PRIOR APPLICATION NUMBER: US 09/104,063
; PRIOR FILING DATE: 1988-06-24
; PRIOR APPLICATION NUMBER: US 08/701,265
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/664,228
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/076,093
; PRIOR FILING DATE: 1993-06-11
; PRIOR APPLICATION NUMBER: US 07/810,782
; PRIOR FILING DATE: 1991-12-19
; PRIOR APPLICATION NUMBER: US 07/677,211
; PRIOR FILING DATE: 1991-03-29
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-666-689-3
```

Query Match 55.1%; Score 933.4; DB 13; Length 1737;
Best Local Similarity 81.8%; Pred. No. 2.2e-236;
Matches 1154; Conservative 0; Mismatches 216; Indels 41; Gaps 5;

```
Qy 1 ATATACACTCTGTAACTACTCTGAAAGCTGGGGCTGGAGACTATGACTCCCAACAG 60
Db 106 ATATACACTCTGTAACTACTCTGAAAGCTGGGGCTGGAGACTATGACTCCCAACAG 165
Qy 61 GAACCCCTCTTCCGGGATGAAACGTCATTTCAATAGGATCTTCTCTGCCACCATCTAC 120
Db 166 GAACCCCTCTTCCGGGATGAAACGTCATTTCAATAGGATCTTCTCTGCCACCATCTAC 225
Qy 121 TTATCATCTTTTGTACTGGCATAGTGGCAATGGATTGGTGGTGGTGGTGGTGGTGGT 180
Db 226 TCCATCATCTTCTTAACTGGCATTTGGGGCAATGGATTGGTGGTGGTGGTGGTGGT 285
Qy 181 CAGAAGAGCTAAAGAGCATGACGCAAGATACCGGCTGCACCTGTGCTAGTGGCTGACCTC 240
Db 286 CAGAAGAGCTAAAGAGCATGACGCAAGATACCGGCTGCACCTGTGCTAGTGGCTGACCTC 345
Qy 241 CTCTTTGTATCAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 346 CTCTTTGTATCAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 405
Qy 301 AAATTTTGTGTAGGCTGTCCATATCATCTATCATCTGTCAACCTGTACAGCAGCGTTCTC 360
Db 406 AACTTCTATGCAAGGAGTCCATGTCTATCATACAGTCAACCTGTACAGCAGTGTCTCTC 465
Qy 361 ATCTGGGCTTTCATCAGCTGAGACCGGTACTCTGCCATTTGTCCAGTGTCCAGCCAA 420
Db 466 ATCTGGGCTTTCATCAGTCTGAGACCGGTACTCTGCCATTTGTCCAGTGTCCAGCCAA 525
Qy 421 AGGCCAAGAACTGCTGGCTGAAAGAGCACTATCTGTGGCGCTGTGGATCCAGCCCTC 480
Db 526 AGGCCAAGAACTGCTGGCTGAAAGAGCACTATCTGTGGCGCTGTGGATCCAGCCCTC 585
Qy 481 CTCTGACTATACCTGACTTTCATCTTTTCCGACGCTCAGCCAGGGGACATCAGTCAAGGG 540
Db 586 CTGCTGACTATCTCCGACTTTCATCTTTTCCGACG-----TCAGTGAAGCA 630
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCTGTGGATGGTGGTGTTC 600
Db 631 GATGACAGATATATCTGTGACCGCTTCTTACCCCAATGACTTGTGGTGGTGTGTCTCCAG 690
```

Search completed: May 17, 2004, 22:14:17
Job time : 840.19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 17, 2004, 05:50:31 ; Search time 4889.52 Seconds
(without alignments)
10345.919 Million cell updates/sec

Title: US-09-367-052-7
Perfect score: 1694
Sequence: 1 atatacactctgataacta.....taaaagtcaatgataaactt 1694

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pla:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	709.4	41.9	1091	13 BX397685	BX397685 BX397685
2	694.8	41.0	1201	13 BX396365	BX396365 BX396365
3	692	40.9	1201	13 BX421287	BX421287 BX421287
4	688.6	40.6	1201	13 BX421511	BX421511 BX421511

5	684	40.4	906	13	BUS23025	AGENCOURT
6	681.4	40.2	1201	13	BX358346	BX358346
7	672.2	39.7	1051	12	BM545259	AGENCOURT
8	666.6	39.4	1201	13	BX353253	BX353253
9	647.6	38.2	1201	13	BX462627	BX462627
10	624.6	36.9	1201	13	BX418530	BX418530
11	621.6	36.7	793	9	AA182270	mt83f01.r
12	621.2	35.7	639	9	AA816049	AA816049
13	608.2	35.9	1201	13	BX462182	vt14h02.r
14	603	35.6	952	13	BX362477	BX362477
15	602.2	35.5	872	14	CD251331	AGENCOURT
16	582.4	34.4	945	13	BX443258	BX443258
17	580.2	34.3	911	13	BX443334	BX443334
18	576	34.0	904	10	BF100790	601753725
19	575.8	34.0	910	12	BT762229	603049139
20	572	33.8	600	13	BUT919690	6033-58 M
21	569	33.6	891	12	BT824663	603033815
22	564	33.3	886	13	BX443228	BX443228
23	558.6	33.0	950	13	BQ718617	AGENCOURT
24	552.6	32.6	765	12	BM387369	UI-R-CN1-
25	551.6	32.6	582	12	BG145042	ut75a01.Y
26	549.6	32.4	749	14	CD466324	LeukON2.3
27	549	32.4	1099	13	BX397684	BX397684
28	547.6	32.3	581	14	CF540844	UI-N-GNO-
29	546.8	32.3	928	14	CD389217	AGENCOURT
30	544.4	32.1	958	12	BG174412	AGENCOURT
31	539.6	31.9	1034	12	BM920800	AGENCOURT
32	539.4	31.8	1201	13	BX358340	BX358340
33	533	31.5	876	14	CA454255	AGENCOURT
34	532.4	31.4	1201	13	BX421141	BX421141
35	529.2	31.2	827	12	BT761664	603046395
36	524.8	31.0	528	12	BG277031	uv10e03.Y
37	524.8	31.0	750	12	BG613352	602641234
38	522	30.8	955	12	BG173867	602333928
39	520.8	30.7	753	12	BG915636	602814106
40	510.6	30.1	585	10	BE527479	BS627479
41	510.2	30.1	875	12	BT756157	uv52b07.Y
42	507	29.9	1201	13	BX418529	BX418529
43	502.4	29.7	976	13	BUI46456	AGENCOURT
44	491.4	29.0	913	12	BM051973	603638877
45	490.4	28.9	782	14	CB985366	AGENCOURT

ALIGNMENTS

RESULT 1
BX397685
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX397685 1091 bp mRNA linear EST 13-MAY-2003
BX397685 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0DI041Y013 5-PRIME, mRNA sequence.

BX397685
BX397685
BX397685.1 GI:30621359
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1091)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI041A070P1&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODI041AH07QPI.

FEATURES

source
Location/Qualifiers
1. .1091
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI041Y013"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 41.9%; Score 709.4; DB 13; Length 1091;
Best Local Similarity 85.4%; Pred. No. 2.4e-106; Indels 16; Gaps 3;
Matches 828; Conservative 2; Mismatches 124;

QY 1 ATATACATCTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAG 60
DB 136 ATATACATCTCAGATAACTTACACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAAG 195
QY 61 GAACCTGCTTCGGGATGAAAAGTCCATTTCAATAGGATCTTCTGCCACCATCTAC 120
DB 196 GAACCTGTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCTGCCACCATCTAC 255
QY 121 TTTCATCATCTTCTGACTGGCATAGTTCGGCAATGATGGTGATCCTGGTATGGGTTAC 180
DB 256 TCCATCATCTTCTTAACCTGCATTTGTGGCATGATGGTATCTCTGTCATGGTTAC 315
QY 181 CAGAGAACTAAGAGCATGACGACAAAGTACCGGCTGCACCTCTCAGTGGTACCTC 240
DB 316 CAGAGAACTAAGAGCATGACGACAAAGTACCGGCTGCACCTCTCAGTGGTACCTC 375
QY 241 CTCCTTGTCTACACTCCCTCTTGGGAGTTGATGCCATGGCTGACTGGTACTTTGGG 300
DB 376 CTCCTTGTCTACACTCCCTCTTGGGAGTTGATGCCATGGCTGACTGGTACTTTGGG 435
QY 301 AAATTTTGTGAAGCTGTCATATCATCTACATCTGTCACCTCTACAGAGGTTCTC 360
DB 436 AACTTCTCTATGCAAGGACGATGCTATCTACACAGTCAACCTCTACAGAGTGTCTC 495
QY 361 ATCCCTGGCTTCATCAGCTGACCGGTACCTCGCCATGTCACGCCACCAACAGTCAA 420
DB 496 ATCCCTGGCTTCATCAGCTGACCGGTACCTCGCCATGTCACGCCACCAACAGTCAA 555
QY 421 AGGCCAAGAAACTGTGGCTGAAAAGGAGTCTATGTGGGCTGTGGATCCAGCCCTC 480
DB 556 AGGCCAAGAAAGCTGTGGCTGAAAAGGAGTCTATGTGGGCTGTGGATCCCTGCCCTC 615
QY 481 CTCCTGACTATACCTGACTTCTCTTTGGCAGCTCAGCCAGGGGAGCATCAGTCAGGGG 540
DB 616 CTGCTGACTATTTCCGACTTCTCTTTGCAACGTCAGTGANGCA----- 661
QY 541 GATGACAGGTACATCTGTGACCGCTTTTACCCCGATAGCTGTGTGATGGTGTTCAC 600
DB 662 GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGTGTTCAG 721
QY 601 TTCCAGCATATATAGTGGGTTCTATCTCTGCCCCGATCGTCATCTCTCTGTACTGTC 660
DB 722 TTTCAGCATATATAGTGGGTTCTATCTCTGCTGTATTTGTCTCTCTCTCTCTATTC 781
QY 661 ATCATCATCTCTAAGCTGTACATCCAGGGCCACAGAGCGCAAGGCCCTCAAGACG 720
DB 782 ATTATCATCTCCAGCTGTACATCTCCAGGGCCACAGAGCGCAAGGCCCTCAAGACC 841
QY 721 ACAGTCATCTCATCTCT-AGCTTTCTTTGCTGCT-GGCTGCCATATTATGTGGGATCA 778
DB 842 ACAGTCATCTCATCTCTGCGCTTCTTCTGCTGTTTGGCTGCTTACTACATGGATCA 901
QY 779 GCATGACCTCTTCATCTTTGGGAGTCATCAAGCAGGATGTGACTTCAGAGCATTTG 838

DB 902 GCATCGACTCTCTTCATCTCTCTGGAATCATCAAGCAAGGGTGTGAGTTTGAGACACTG 961
QY 839 TGCACAAGTGGATCTCCATCAGAGAGCGCTCGCTTCTTCACTGCTTGGCTGAACCCCA 898
DB 962 TGCACAAGTGGATTTCCATCAGAGAGCGCTAGCTTTTTCACATGTTTGAACCCCA 1021
QY 899 TCTCTATGCTTCTCTCGGGGCAAGTTCAAAAGCTCTGCCAGCATGCACTCAATCCA 958
DB 1022 TCTCTATGCTTCTCTGGAGCCAAATTTAAACCTCTGCCAGCAGCACTCACTCTG 1081
QY 959 TGAGCAGAGG 968
DB 1082 TGAGCAGAGG 1091

RESULT 2
BX396365

LOCUS 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX396365 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI015YL20 5-PRIME, mRNA sequence.

ACCESSION

VERSION BX396365

KEYWORDS

SOURCE EST.

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li, W. B., Gruber, C., Jessee, J., and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1272.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cg1-bin/cluster.cgi?seq=CSODI015DF100P1&cluster=1272.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODI015DF100P1.

FEATURES

source

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODI015YL20"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 41.0%; Score 694.8; DB 13; Length 1201;

Best Local Similarity 83.9%; Pred. No. 5.9e-104;

Matches 842; Conservative 7; Mismatches 134; Indels 21; Gaps 5;

QY 1 ATATACATCTCTGATAACTACTCTGAAGAAGTGGGGTCTGGAGACTATGACTCCAAACAG 60

DB 134 ATATACATCTCAGATAACTTACACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAAG 193

QY 61 GAACCTGCTTCGGGATGAAAAGTCCATTTCAATAGGATCTTCTGCCACCATCTAC 120

DB 194 GAACCTGTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTCTGCCACCATCTAC 253

QY 121 TTTCATCATCTTCTGACTGGCATAGTGGCAATGGATTGGTGTATCTCTGCTGCGGTTAC 180

DB 254 TCCATCATCTTCTTAACCTGTCATTTGGGCAATGGATTGGTGTATCTCTGCTGCGGTTAC 313

QY 181 CAGAAGAGCTAAGAGCATGACGGACAGATGACGGCTGACCTGTGAGTGGCTGACCTC 240
DB 314 CAGAAGAACTGAGAAGCATGACGGACAGATGACGGCTGACCTGTGAGTGGCGACCTC 373
QY 241 CTCTTTGTCATCAGCTCCCTCTCTGGGAGTTCATGCGATGGTGTGACTGTACTTTGGG 300
DB 374 CTCTTTGTCATCAGCTCCCTCTCTGGGAGTTCATGCGATGGTGTGACTGTACTTTGGG 433
QY 301 AAATTTTGTATAGGCTGTCCATATATATATATATATATATATATATATATATATAT 360
DB 434 AACTTCTTATGCAAGGAGTCCATGTATATATATATATATATATATATATATATATAT 493
QY 361 ATCTGGGCTTCATCAGCTGAGCGGTACCTGCGCATGTCACGCGCCACCAACAGTCAA 420
DB 494 ATCTGGGCTTCATCAGCTGAGCGGTACCTGCGCATGTCACGCGCCACCAACAGTCAA 553
QY 421 AGGCCAAGAACTGCTGGCTGAAAGGACAGTCTATGTGGCGCTGTGATCCAGGCCCTC 480
DB 554 AGGCCAAGAACTGCTGGCTGAAAGGAGTCTATGTGGCGCTGTGATCCAGGCCCTC 613
QY 481 CTCTGACTATACCTGACTTCATCTTTGCGAGCTGACCTGCGCATGTCACGCGCCACCA 540
DB 614 CTGCTGACTATTCGCGACTTCATCTTTGCCAAG-----TCAGTAGGGCA 658
QY 541 GATCAGAGTACATCTGTGAGCGGCTTTACCCCGATAGCTGTGAGTGGTGTGTTCAA 600
DB 659 GATCAGAGTATATCTGTGAGCGGCTTCATCCCAATGACTTGTGGTGTGTTCCAG 718
QY 601 TTCAGCATATATAGTGGTGTCTATCTGCGCGGACATGTCATCTCTCTGTACTGC 660
DB 719 TTTACGACATCATGTTGGCTTTATCTGCTGCTGATTTGTCATCTCTGCTGTATTC 778
QY 661 ATCATCATCTTAAGCTGTACACTTCCAAAGGCGCACAGAGGCGCAAGGCGCTCAAGACG 720
DB 779 ATTATCATCTCCAGCTGTACACTTCCAAAGGCGCACAGAGGCGCAAGGCGCTCAAGAC 838
QY 721 ACAGTCATCTCATCTGCTTTGGAGTTCATCAAGCAAGGATGAGTTCGAGAGCATGTG 840
DB 839 ACAGTCATCTCATCTGCTTTGGAGTTCATCAAGCAAGGATGAGTTCGAGAGCATGTG 898
QY 781 ATCGACTCTTCATCTGCTTTGGAGTTCATCAAGCAAGGATGAGTTCGAGAGCATGTG 840
DB 899 ATCGACTCTTCATCTGCTTTGGAGTTCATCAAGCAAGGATGAGTTCGAGAGCATGTG 958
QY 841 CACAAGTGGATTCATCAGAGAGGCTTCGCTTTTCCACTGTGCTGAAACCCCATC 900
DB 959 CACAAGTGGAT-TYCATCAGCGAGGCGCTAGCTTTTCCACTGTGCTGAA-CCCATC 1016
QY 901 CTCTATGCTTCTCGGGGCAAGTTCAAAGCTCTGCGGAGCATGCACTCAACTCCATG 960
DB 1017 CTCTATGCTTCTCGGGGCAAGTTCAAAGCTTCTTCCACTGTGCTGAA-CCCATC 1073
QY 961 AGCAGAGGCTCCAGGCTCAAGATCCTTTCCAAAGGAAAGCGGG 1004
DB 1074 AGCAGAGGCKWAG-CTCAGATCTCTCWAAGGAGCGGG 1116

RESULT 3
BX421287
LOCUS
DEFINITION BX421287 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
clone CS0DG004YG09 5-PRIME, mRNA sequence.
ACCESSION BX421287
VERSION
KEYWORDS
SOURCE EST. GI:30766068
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DG004YD05QPl&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DG004YD05QPl.

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DG004YG09"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 40.9%; Score 692; DB 13; Length 1201;
Best Local Similarity 76.6%; Pred. No. 1.7e-103;
Matches 847; Conservative 43; Mismatches 193; Indels 23; Gaps 4;
QY 1 ATATACACTTCTGATAACTACTCTGAGAAAGTGGGTCTGGAGACTATGACTCCAAACAAG 60
DB 118 ATATACACTTCTGATAACTACTCTGAGAAAGTGGGTCTGGAGACTATGACTCCAAACAAG 177
QY 61 GAACCTGCTTCGGGATGAAAGCTTCATTTCAATAGGATCTTCTGCGCCACCACTAC 120
DB 178 GAACCTGCTTCGGGATGAAAGCTTCATTTCAATAGGATCTTCTGCGCCACCACTAC 237
QY 121 TTCTATCATCTCTTGTGATGGCATAGTCGGCAATGATTTGGTATCTGCTGTGATGGGTTAC 180
DB 238 TCCTATCATCTCTTGTGATGGCATAGTCGGCAATGATTTGGTATCTGCTGTGATGGGTTAC 297
QY 181 CAGAAGAACTGAGAAGCATGACGGACAGTACAGGCTGCACCTGTGCTGAGTGGCTGACCTC 240
DB 298 CAGAAGAACTGAGAAGCATGACGGACAGTACAGGCTGCACCTGTGCTGAGTGGCTGACCTC 357
QY 241 CTCTTTGTCATCAGCTCCCTCTCTGGGAGTTCATGCGATGGTGTGACTGTACTTTGGG 300
DB 358 CTCTTTGTCATCAGCTCCCTCTCTGGGAGTTCATGCGATGGTGTGACTGTACTTTGGG 417
QY 301 AAATTTTGTATAGGCTGTCCATATATATATATATATATATATATATATATATATAT 360
DB 418 AACTTCTTATGCAAGGAGTCCATGTATATATATATATATATATATATATATATATAT 477
QY 361 ATCTGGGCTTCATCAGCTGAGCGGTACCTGCGCATGTCACGCGCCACCAACAGTCAA 420
DB 478 ATCTGGGCTTCATCAGCTGAGCGGTACCTGCGCATGTCACGCGCCACCAACAGTCAA 537
QY 421 AGGCCAAGAACTGCTGGCTGAAAGGACAGTCTATGTGGCGGTGTGATCCAGGCCCTC 480
DB 538 AGGCCAAGAACTGCTGGCTGAAAGGAGTGTGATTTGGGCTGTGATCCCTGCCCTC 597
QY 481 CTCTGACTATACCTGACTTCATCTTTGCGAGCTGACGCGCGGAGACATCATAGTAGGGG 540
DB 598 CTCTGACTATTTCCGACTTCATCTTTTGGCAAG-----TCAGTAGGGCA 642
QY 541 GATGAGAGTATATCTGTCAGCGGCTTTTACCCGATAGCTGTGATGGTGTGTTCAA 600
DB 643 GATGAGAGTATATCTGTGACCGCTTCTACCCCAATGATTTGGGAGTGTGTTCCAG 702
QY 601 TTCAGCATATATAGTGGTGTCTCATCTGCGCGGATCGTCACTCTCTCTGTACTGC 660


```
Db 1027 CTCATGCTTCTCTTGGAGCAAAATTTAAACTCT 1060
RESULT 5
BU523025 906 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT 10154335 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6529663 5', mRNA sequence.
ACCESSION BU523025
VERSION BU523025.1 GI:22833475
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 906)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM4130 row: c column: 07
High quality sequence stop: 668.
FEATURES
source
1. 906
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529663"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP Co24"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 40.4%; Score 684; DB 13; Length 906;
Best Local Similarity 98.1%; Pred. No. 3.6e-102;
Matches 714; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
Qy 1 ATATACACTTCTGATTAACACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCAAACAG 60
Db 84 ATATACACTTCTGATTAACACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCAAACAG 143
Qy 61 GAACCTCTGCTCCGGGATGAACGTCCTCAATAGGATCTTCTGCGCCACCATCTAC 120
Db 144 GAACCTCTGCTCCGGGATGAACGTCCTCAATAGGATCTTCTGCGCCACCATCTAC 203
Qy 121 TTATCATCTTCTTGATGCGCATAGTCGGCAATGGAATGGTGTGATCTCTGTCATGGGTTAC 180
Db 204 TTATCATCTTCTTGATGCGCATAGTCGGCAATGGAATGGTGTGATCTCTGTCATGGGTTAC 263
Qy 181 CAGAGAGCTAGGAGCATGACGACAGTACCGCTGACCTGTGACGTGACCTGACCTC 240
Db 264 CAGAGAGCTAGGAGCATGACGACAGTACCGCTGACCTGTGACGTGACCTGACCTC 323
Qy 241 CTCCTTTGTCATACACTCCCTCTTCTGGCAGTGTGATGCCATGGCTGACTGTGTTGG 300
Db 324 CTCCTTTGTCATACACTCCCTCTTCTGGCAGTGTGATGCCATGGCTGACTGTGTTGG 383
Qy 301 AAATTTTGTGAAGCTGTCCATATCACTACATGTCACCTCTACAGACGGTTCTC 360
Db 384 AAATTTTGTGAAGCTGTCCATATCACTACATGTCACCTCTACAGACGGTTCTC 443
```

```
Qy 361 ATCTTGGCTTCATCAGCTGGACCGGTACCTCGCATTTGTCACGCCAACAGTCAA 420
Db 444 ATCTTGGCTTCATCAGCTGGACCGGTACCTCGCATTTGTCACGCCAACAGTCAA 503
Qy 421 AGCCCAAGGAACACTGCTGGCTGAAAGGACGTCATATGTTGGGCTGTGATCCACCCCTC 480
Db 504 AGCCCAAGGAACACTGCTGGCTGAAAGGACGTCATATGTTGGGCTGTGATCCACCCCTC 563
Qy 481 CTCCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCCAGGGGACATCAGTCAGGG 540
Db 564 CTCCTGACTATACCTGACTTTCATCTTTCCGACGTCAGCCAGGGGACATCAGTCAGGG 623
Qy 541 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCCTGTGATGGTGGTGTTCAA 600
Db 624 GATGACAGGTACATCTGTGACCGCTTTACCCCGATAGCCTGTGATGGTGGTGTTCAA 683
Qy 601 TTCCAGCATATAATGGT-GGGTCTCATCTCTGCGCGGACATCGTCATCTCTCTCTTACTG 659
Db 684 TTCCAGCATATAATGGTGGGGTCTCATCTCTGCGCGGACATCGTCATCTCTCTCTTACTG 743
Qy 660 GATGATCATCTCTAAGCTGTACATCTCAAGGG---CCACAGAGAGCGCAAGGCCCTCAA 716
Db 744 CATCATCATCTCTAAGCTGTACATCTCAAGGGGCCACCCAGAGAACGCCAGGCCCTCAA 803
Qy 717 GAGGACAG 724
Db 804 GAGGACAG 811
RESULT 6
BX358346 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX358346
DEFINITION BX358346 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1036YJ17 5-PRIME, mRNA sequence.
ACCESSION BX358346
VERSION BX358346.1 GI:30370215
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1036C509Qp1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036YJ17"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 40.2%; Score 681.4; DB 13; Length 1201;
Best Local Similarity 85.3%; Pred. No. 9.1e-102;
Matches 783; Conservative 4; Mismatches 116; Indels 15; Gaps 2;
Qy 1 ATATACACTTCTGATTAACACTCTGAGAGAGTGGGTCTGGAGACTATGACTCCAAACAG 60
```


481 CTCCTGACTATACCTGACTCTCTTTGGGAGCTCAGCAGGGGAGCATCAGTCAGGGG 540
 Db CTGCTGACTATTCCTGAGCTTCTATCTTTGCCAAG- - - - -TCAGTGAGGCA 607
 541 GATCAGAGTACATCTGTGACCGCTTACCCGATAGCTGTGGATGGTGGTTCAC 600
 Db GATCAGAGATATCTGTGACCGCTTCTACCCCAATGACTGTGGGTGGTGTCCAG 667
 601 TTCAGGATATAAATGGTGGGCTCTCATCTGCGCGGATCGTCTATCTCTCTGTACTGC 660
 Db TTTGAGCAGATCATGTTGGGCTTATCTGCTGCTATTTGTCATCTCTGCTATTGC 727
 661 ATCATCATCTTAAGCTGTGACATCTCAGAGGCGCACAGAGCGCAAGGCGCTCAGAGC 720
 Db ATTATCATCTCAAGCTGTACATCTCAGAGGCGCACAGAGCGCAAGGCGCTCAGAGC 787
 721 ACAGTCATCTCATCTAGCTTTCTTTGGCTGTGGCTGCGCATATTATGTGGGATCAGC 780
 Db ACAGTCATCTCATCTAGCTTTCTTTGGCTGTGGCTGCGCATATTATGTGGGATCAGC 847
 781 ATGACTCTCTTATCTTTGGGAGTCAATCAAGCAA-GAGTGTGACTTTCAGAGCAATGT 839
 Db ATGACTCTCTTATCTTCTGGAATCATCAAGCAAGGGGTGTGAGTTTGAGAACGGT 907
 840 GCACAGTGTATCTCCATCAGAG-GGCTCTGCTTCTCCACTGTGCTGAAACCCCA 898
 Db GCACAGTGTATCTCCATCAGAGGCGCTTCTTCTTCACTGTGCTGAAACCCCA 967
 899 TCCTCTATGCTT--CCTGGGGGCAAGTTCAAGAGCTCTGCCAGATGCACTCACTC 956
 Db TCCTCTATGCTTCTTCTGGAGCGCAATTTAAACCTTGTGCCAGACGCACTCCACTC 1027
 957 CATG-AGCAGAGGCTCCAGCCTC 978
 Db TGTGAACAAGGGTCCAGCCTC 1050

RESULT 8
 BX353253
 LOCUS
 DEFINITION BX353253 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC003YG03 5-PRIME, mRNA sequence.
 ACCESSION BX353253
 VERSION BX353253.1 GI:30347726
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1272.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC003AD02QP1&cluster=1272.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC003AD02QP1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC003YG03"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

FEATURES

source

ORIGIN
 Query Match 39.4%; Score 666.6; DB 13; Length 1201;
 Best Local Similarity 82.6%; Pred. No. 2.4e-99;
 Matches 808; Conservative 6; Mismatches 146; Indels 18; Gaps 4;
 QY 1 ATATACATCTTCGATACCTCTCAAGAAGTGGGCTGTGGAGACTATGACTCCAAACAG 60
 Db 151 ATATACATCTTCGATACCTCTCAAGAAGTGGGCTGTGGAGACTATGACTCCAAACAG 210
 QY 61 GAACCTCTCTTCGGGATCAAAACCTCCATTTCAATAGGATCTTCTGCCCCACCACTAC 120
 Db 211 GAACCTCTCTTCGGGATCAAAACCTCCATTTCAATAGGATCTTCTGCCCCACCACTAC 270
 QY 121 TTTCATCATCTTCTTCACTGCGGATAGTCGGGATGATGATGATGATGATGATGATGAT 180
 Db 271 TCCATCATCTTCTTCACTGCGGATAGTCGGGATGATGATGATGATGATGATGATGATGAT 330
 QY 181 CAGAAAGAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTCTCAGTGGCTGACCTC 240
 Db 331 CAGAAAGAGCTAAGGAGCATGACGACAAAGTACCGGCTGCACCTCTCAGTGGCTGACCTC 390
 QY 241 CTCTTGTGATCACACTCCCTTCTGGGAGTGTGATGATGATGATGATGATGATGATGAT 300
 Db 391 CTCTTGTGATCACACTCCCTTCTGGGAGTGTGATGATGATGATGATGATGATGATGAT 450
 QY 301 AAATTTTGTGTAAAGCTGTCCATATCATCTACACTGTCAACCTCTACAGACAGCGTTCTC 360
 Db 451 AAATTTTGTGTAAAGCTGTCCATATCATCTACACTGTCAACCTCTACAGACAGCGTTCTC 510
 QY 361 ATCTGGGCTTCATCAGGCTGAGCCGCTACCTGCGCATTTGCCAGCAACCAAGTCAAA 420
 Db 511 ATCTGGGCTTCATCAGGCTGAGCCGCTACCTGCGCATTTGCCAGCAACCAAGTCAAA 570
 QY 421 AGGCCAAGGAAGTCTGCTGGTGAAGGAGCTATGTTGGGCTGTGATGCCAGCCCTC 480
 Db 571 AGGCCAAGGAAGTCTGCTGGTGAAGGAGCTATGTTGGGCTGTGATGCCAGCCCTC 630
 QY 481 CTCTGACTATACCTGACTTCACTCTTCCGAGCTCAGCCAGGGGAGACATCAGTCAGGG 540
 Db 631 CTCTGACTATACCTGACTTCACTCTTCCGAGCTCAGCCAGGGGAGACATCAGTCAGGG 675
 QY 541 GATGACAGTATACCTGAGCCGCTTACCCCGATAGCTGTGATGATGATGATGATGATGAT 600
 Db 676 GATGACAGTATACCTGAGCCGCTTACCCCGATAGCTGTGATGATGATGATGATGATGAT 735
 QY 601 TTTCAGCATATTAATGTTGGTCTCATCTGCGCGGATCGTCACTCTCTCTGTTACTGC 660
 Db 736 TTTCAGCATATTAATGTTGGTCTCATCTGCGCGGATCGTCACTCTCTCTGTTACTGC 795
 QY 661 ATCATCATCTTAAAGCTGTCACTTCCAGGGGCAACCAAGAGCGCAAGGCTTCAAGAG 720
 Db 796 ATATCATCTTCAAGCTGTCACTTCCAGGGGCAACCAAGAGCGCAAGGCTTCAAGAG 855
 QY 721 ACAGTCATCTCTCATCTGACTTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 856 ACAGTCATCTCTCATCTGACTTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
 QY 781 ATGACTCTCTTCACTCTTGGAGTCAATCAAGAGGAGTGTGATGATGATGATGATGATGAT 840
 Db 916 ATGACTCTCTTCACTCTTGGAGTCAATCAAGAGGAGTGTGATGATGATGATGATGATGAT 975
 QY 841 CACAGTGGATCTCCATCAGAGGCGCTCCGCTTCTTCCAGTGTGCTGCTGCTGCTGCTGCT 900
 Db 976 CACAGTGGATCTCCATCAGAGGCGCTCCGCTTCTTCCAGTGTGCTGCTGCTGCTGCTGCT 1033
 QY 901 CTCTATGCTTCTCTGGGGCCCAAGTTCAAAAGCTCTGCCAGGAGTCACTCAACTCCATG 960
 Db 1034 CTCTATG-CTTCTTGGAGCAAAATTAAGTCTGCGAGCAGCACTCACTCTCTGTTAGAGA 1092

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

QY 961 AGCAGAGGCTCCAGCCTC 978
Db 1093 GGGTCMRCTCAGATCTC 1110

RESULT 9
LOCUS BX462627 1201 bp mRNA linear EST 22-MAY-2003
DEFINITION BX462627 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CS0DH005YK10 5-PRIME, mRNA sequence.
ACCESSION BX462627
VERSION BX462627.1 GI:31029445
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1201)
TITLE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DH005YK10&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DH005BF05Qp1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH005YK10"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 38.2%; Score 647.6; DB 13; Length 1201;
Best Local Similarity 84.8%; Pred. No. 3,1e-96;
Matches 767; Conservative 0; Mismatches 119; Indels 18; Gaps 3;

QY 1 ATATACACTTCTGATACTACTCTGAAGAAGTGGGGTCTGAGACTATGACTCCACAAG 60
Db 148 ATATACACTTCAGATNACTACACGAGGAATGGGCTCAGGGACTATGACTCCTGAAG 207
QY 61 GAACCCCTGTTCCGGGATGAACAGTCCATTTCAATAGATTTCTTCGCCCAACATCTAC 120
Db 208 GAACCCCTGTTCCGGTGAAGAAATGCTAAATTTCAATAAATCTTCTGCCCAACATCTAC 267
QY 121 TTCTATCATCTTCTGACTGTCATAGTCGCAATGGATTGGTATCTCTGGTCACTGGGTTAC 180
Db 268 TCCATCATCTTCTTAATCGGCATTTGGGCATATGGATTGGTATCTCTGGTCACTGGGTTAC 327
QY 181 CAGAAGAGCTAAGGAGCATGACGACAGTACCGCTGCACCTGTGCAGTGGCTGACCTC 240
Db 328 CAGAAGAACTGAGAAGCATGACGACAGTACAGGCTGCACCTGTGCAGTGGCGGACCTC 387
QY 241 CTCCTTGTGCATCACACTCCCTCTTGGCAGTTATGCGATGGCTGACTGGTACTTTGGG 300
Db 388 CTCCTTGTGCATCACGCTTCCCTCTTGGCAGTTATGCGGTCAGTGGCAACTGGTACTTTGGG 447

```

```

QY 301 AAATTTTGTGTAAGGCTGTCCATATCATCTACACTGTCAACCTCTCAGCAGCGCTTCTC 360
Db 448 AACTTCTCTATGCAAGGCAAGTCCATGTCTATACACAGTCAACCTCTCAGCAGTGTCTCTC 507
QY 361 ATCTTGGGCTTTCATCAGCTGGAACGGTACCTCGCCATGTCCAGCCCAACAGTCAA 420
Db 508 ATCTTGGGCTTTCATCAGTCTGGAACGGTACCTCGCCATGTCCAGCCCAACAGTCAAG 567
QY 421 AGGCCAAGAAACTCTGCTGGCTGAAAGGCAAGTCTATGTGGGCGTCTGGATCCAGCCCTC 480
Db 568 AGGCCAAGGAAGCTGTGGCTGAAAGGCTGTCTATGTGGGCGTCTGGATCCCTGCCCCCTC 627
QY 481 CTCCTGACTATACCTGACTTTCATCTTTCGCCAGCTCAGCCAGGGGGACATCAGTCAGGGG 540
Db 628 CTGCTGACTATTCCTGCACTTTCATCTTTCGCCAAGC-----TCAGTCAGGCA 672
QY 541 GATGACAGGTATACCTGTGACCGCTTTACCCCGATAGCTGTGATGGTGGTGTTCAA 600
Db 673 GATGACAGATATATCTGTGACCGCTTTCACCCCATGATCTTGGGTGGTGTGTTCAG 732
QY 601 TTCCAGCATATAATGGTGGTCTCATCTGCCGGGATCGTCATCTCTCTGTACTTGC 660
Db 733 TTTTCAGCATCATGTTGGTGGCTTATCTCTGCTGGTATTTGTCTCTCTCTCTATTGC 792
QY 661 ATCATCATCTCTAAGCTGTCACTCCCAAGGGCCACCAAGAGCGCAAGCCCTCAAGACG 720
Db 793 ATATCATCTCTCAAGCTGTCACTCCCAAGGGCCACCAAGAGCGCAAGCCCTCAAGACC 852
QY 721 ACAGTCACTCCATCCTAGCTTTCTTCCCTGCTGGTGGCCATATATGTGGGATCAGC 780
Db 853 ACAGTCACTCCATCCTAGCTTTCTTCCCTGCTGGTGGCCATATATGTGGGATCAGC 912
QY 781 ATCGACTCTTCATCCTTTTGGGAGTGTCAAGCAAGGATGTGACTTCGAGAGCAATTGTG 840
Db 913 ATCGACTCTTCATCCTCTCTGAAATCATCAAGCAAGGTTGTGAGTTTGAGAACACTGTG 972
QY 841 CACAAGTGGATCTCCATCAGAGGCCCTCGCCTTCTTCCACTGTGTGCTGAACCCCATC 900
Db 973 CACAAGTGGAT-TTCATCAGCAGGACC--TAGCTTTCTTCTACTGTGTGCTGAACCCCATC 1029
QY 901 CTCT 904
Db 1030 CTAT 1033

RESULT 10
LOCUS BX418530 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX418530 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF006YJ16 5-PRIME, mRNA sequence.
ACCESSION BX418530
VERSION BX418530.1 GI:30658606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1201)
TITLE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1272.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF006YJ16&cluster=1272.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

```


QY 493 CCTGACTTCATCTTTGCGGACGTCAGCCAGGGGAGCATCAGTCAGGGGGATGACAGGTAC 552
 Db 1 CTTGACTTCATCTTTGCGGACGTCAGCCAGGGGAGCATCAGTCAGGGGGATGACAGGTAC 60
 QY 553 ATCTGTGACCGCTTTACCCGATAGACCTGTGGATGGTGGTGTTCATATTCAGCATATA 612
 Db 61 ATCTGTGACCGCTTTACCCGATAGACCTGTGGATGGTGGTGTTCATATTCAGCATATA 120
 QY 613 ATGGTGGGTCTCATCTGCGCGGCATCGTCATCTCTCTCTGTACTGCATCATCATCTCT 672
 Db 121 ATGGTGGGTCTCTGCTGCGCGGCATCGTCATCTCTCTCTGTACTGCATCATCATCTCT 180
 QY 673 AGCTGTGACACTCCAGGGGCCACAGAGGCCAGAGGCCCTCAAGACAGCATCATCTCT 732
 Db 181 AGCTGTGACACTCCAGGGGCCACAGAGGCCAGAGGCCCTCAAGACAGCATCATCTCT 239
 QY 733 ATCTAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
 Db 240 ATCTAGCTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
 QY 793 ATCTTTTGGGAGTCATCAAGCAAGATGTGACTTCGAGAGCATGTGTCACAAATGGATC 852
 Db 300 ATCTTTTGGGAGTCATCAAGCAAGATGTGACTTCGAGAGCATGTGTCACAAATGGATC 359
 QY 853 TCATCACAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
 Db 360 TCATCACAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
 QY 913 CTCGGGGCAAGTTCAAAGCTCTCCGAGCATGCACTCAACTCCATGAGCAGAGGCTCC 972
 Db 420 CTCGGGGCAAGTTCAAAGCTCTCCGAGCATGCACTCAACTCCATGAGCAGAGGCTCC 479
 QY 973 AGCTTCAGATCTTCCAAAGGAAGGGGTGACACTCTTCGCTTCCAGGAGTCA 1032
 Db 480 AG-CTCAGATCTTTC-CAAGGAAGGGGTGACACTCTTCGCTTCCAGGAGTCA 537
 QY 1033 GAATCTCCAGTTTTCACCTCAGCTAACCCCTTATGCAAGACTTATATA---ATATATA 1088
 Db 538 GAATCTCCAGTTTTCACCTCAGCTAACCCCTTATGCAAGACTTATATA---ATATATA 597
 QY 1089 TATATATATGATAAGAACTTTTATGTTACACATTTTCAGATATAAGAGATGACCA 1148
 Db 598 TATATATATGATAAGAACTTTTATGTTACACATTTTCAGATATAAGAGATGACCA 657
 QY 1149 GTCTGTGACAGTTTCTTTTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208
 Db 658 AGCTGTGACAGTTTCTTTTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711
 QY 1209 TTGTGAGTTTTCATTAATATATAATAATATGTTTTTTTGTGTTTCTGATGTA 1262
 Db 712 TTGTGAGTTTTCATTAATATATAATAATATGTTTTTTTGTGTTTCTGATGTA 765
 RESULT 12
 AA816049
 LOCUS
 DEFINITION
 v14h02.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
 IMAGE:1120659 5', similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
 RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.
 AA816049
 ACCESSION
 AA816049.1 GI:2885645
 VERSION
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 639)
 REFERENCE
 AUTHORS
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Gessel,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE JOURNAL COMMENT

The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:609395
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 510.

FEATURES source

1. 639
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C3H"
 /db_xref="taxon:10090"
 /clone="IMAGE:1120659"
 /cell_line="C2C12"
 /lab_host="DH10B"
 /clone_lib="Barstead mouse myotubes MPLRB5"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTAGAAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [AATTCGATCCTTG], digested with Not I and cloned into the
 Not I and Eco RI sites of the modified p773 vector.
 Library constructed by Bob Barstead. The C2C12 cell line
 (available from ATCC, catalog # CRM-1772) differentiates
 rapidly, forming contractile myotubes and producing
 characteristic muscle proteins."

ORIGIN

Query Match 36.7%; Score 621.2; DB 9; Length 639;
 Best Local Similarity 99.5%; Pred. No. 7,4e-92;
 Matches 623; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 548 GGTACATCTGTGACCGCTTTTACCCGATAGCTGTGGATGGTGGTTCATATCCAGC 607
 Db 14 GGTACATCTGTGACCGCTTTTACCCGATAGCTGTGGATGGTGGTTCATATCCAGC 73
 QY 608 ATATATGCTGGGTCTCATCTGCCCGCATGCTCATCTCTCTCTGTTACTGCATCATCA 667
 Db 74 ATATATGCTGGGTCTCATCTGCCCGCATGCTCATCTCTCTCTGTTACTGCATCATCA 133
 QY 668 TCTCTAAGCTGTTCACACTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGAGCAGTCA 727
 Db 134 TCTCTAAGCTGTTCACACTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGAGCAGTCA 193
 QY 728 TCTCATCTAGCTTTCTTTGCTGCTGCTGCCATATTATGTGGGATCAGATCGACT 787
 Db 194 TCTCATCTAGCTTTCTTTGCTGCTGCTGCCATATTATGTGGGATCAGATCGACT 253
 QY 788 CTTTCATCTCTTTGGGAGTCATCAAGAGGATGTGACTTTCAGAGCATTGTGCACAAGT 847
 Db 254 CTTTCATCTCTTTGGGAGTCATCAAGAGGATGTGACTTTCAGAGCATTGTGCACAAGT 313
 QY 848 GATCTCATATCAGAGGCCCTTCCTTCTTCCATCTGTTGCTGAAACCCCATCTCTATG 907
 Db 314 GGATCTCCATCAGAGGCCCTTCCTTCTTCCATCTGTTGCTGAAACCCCATCTCTATG 373
 QY 908 CTTTCTCGGGGCCAAGTTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAG 967
 Db 374 CTTTCTCGGGGCCAAGTTTCAAAAGCTCTGCCAGCATGCACTCAACTCCATGAGCAGAG 433
 QY 968 GTTCCAGCTCAAGATCTTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTTCCAGG 1027
 Db 434 GCTCCAGCTCGAAGATCTTTTCCAAAGGAAGCGGGGTGGACACTCTTCCGTTCCAGG 493

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1272.r For more information about this cluster, see <http://www.genoscope.cns.tr/cgi-bin/cluster.cgi?seq=CSD0J015CH05QP1&cluster=1272.r>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSD0J015CH05QP1.

FEATURES
source

```

1. 952
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSGDU01SP09"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match	35.6%	Score 603	DB 13	Length 952
Best Local Similarity	85.2%	Pred. No. 6.4e-89		
Matches 685	Conservative 4	Mismatches 100	Indels 15	Gaps 1
QY	1	ATATACACTCTGTAAGTACTCTGTAAGAGATGGCGCTCGGAGATCATGATCCCAAG	60	
Db	154	ATATACACTTCAGTAACTATACCGAGGAATGGGCTCAGGGATATGATCCATGAAG	213	
QY	61	GAACCCCTGCTTCCGGGATGAAAAGTCCTATTTCAATPAGGATCTTCTCCCAACCATCTAC	120	
Db	214	GAACCCCTGTTTCCGTCGAGAAAATGCTAATTTCAATAAAATCTTCTCTGCCACCATCTAC	273	
QY	121	TTTCATCATCTTCTTGACTGTCATAGTCGGCAATGGATTCCTGATCCTGGTCATGGGTAC	180	
Db	274	TCCATCATCTTCTTAACTGGCATTTGGGCATATGGTATCTCTGGTCATGGGTAC	333	
QY	181	CAGAAGAAGCTAAGGAGCATGACGGACAAGTACCGGCTCCACTCTGTCAGTGGCTGACCTC	240	
Db	334	CAGAAGAAGCTGAGAAGCATGACGGACAAGTACAGGCTGCACCTGTCAGTGGCGGACCTC	393	
QY	241	CTCTTTTGTCATCACATCCCTCTTGGGCAGTTGATGCCATGGCTGATGCTGACTTTGGG	300	
Db	394	CTCTTTTGTCATCAGCTTCCCTTCTGGGCAGTTGATGCCGTGGCAACTGGTACTTTGGG	453	
QY	301	AAATTTTGTGTAAAGGCTGTCCATATCATCTACACTGTCAACTCTACAGCAGGCTTCTC	360	
Db	454	AACCTCTCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACTCTACGGCAGTGTCTC	513	
QY	361	ATCTCGGCTTCATCAGCTTGGACGGTACCTCGCAATGTCACGCCACCAACAGTCAA	420	
Db	514	ATCTCGGCTTCATCAGTCTGACCGCTACTCTGCCATTTGTCCAGCCACCAACAGTCAAG	573	
QY	421	AGGCCAAGGAACTGCTGGCTGAAAAAGGCAGTCTATGTGGCGCTCTGATCCCAAGCCCTC	480	
Db	574	AGGCCAAGGAACTGTTGGCTGTAAAGGTGGTCTATGTTGGGCTCTGATCCCTGCCCTC	633	
QY	481	CTCTGACTATACCTGACTTCATCTTTTGGCAGCTCAGCCAGGGGGACATCAGTCAGGGG	540	
Db	634	CTCTGACTATTTCCCGACTTCATCTTTGGCAACG-----TCAGTANGGCA	678	
QY	541	GATGACAGGTACATCTGTGACCGCCTTTTACCCCGATAGCCTGTGATGGTGGTGTTCAA	600	
Db	679	GATGACAGATATATCTGTGACCGCTCTTACCCCAATGACTGTGTGGTGTGTTCAG	738	
QY	601	TTCCAGCATATAATGGTGGGTCTCATCTCTGCCCGGATGTCATCTCTCTGTATTCTGC	660	
Db	739	TTTCAGACATCATGTTGGTGGCCCTATCTCTGCTGGTATTTGTCATCTCTCTGTATTTC	798	
QY	661	ATCATCATCTCTAAGCTGTGCACACTCCAAAGGGCCACCAAGAGCGCAAGGCCCTCAAGC	720	

Db	799	ATTATCATCTCCAGCTGTACACTCAAGGCCACGAAAGCCCTTCAAGACC	858
Qy	721	ACAGTCATCTCATCTCTAGCTTTCTTTGGCTGTGCTGCCATATTATGTGGGATCAGC	780
Db	859	ACAGTCATCTCATCTGCTTTCTTCGCTGTGCTGCTTACTACATTTGGGATCAGC	918
Qy	781	ATCGACTCCTTTCATCTCTTTTGGCA	804
Db	919	ATCGACTCCTTTCATCTCTCTGGAA	942

RESULT 15

[illegible]

ACCESSION	CD251331
VERSION	CD251331.1
KEYWORDS	GI:31011797
EST.	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

ORGANISM

REFERENCE
1 (bases 1 to 872)
AUTHORS
NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: NDAM449 row: h column: 11
High quality sequence stop: 654.

FEATURES

```

1. 872
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30384274"
/tissue_type="Pituitary"
/lab_host="MDH10B-Ton_A (T1 and T5 phase resistances)"
/clone_lib="NIH_MGC_479"
/note="Organ: Brain; Vector: pCMV-Sport6.1; Site_1: EcoRV
(destroyed); Site_2: NotI; Library is oligo-dr primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.1 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
library"

```

ORIGIN

	Query Match	35.5%;	Score 602.2;	DB 14;	Length 872;
	Best Local Similarity	85.9%;	Pred. No. 8.9e-89;		
	Matches 697;	Conservative 0;	Mismatches 98;	Indels 16;	Gaps 2;
QY	159	GGTGATCCTGGTCATGGGTTACCAAGAAGAGCTAAGGAGCATGACGGACAAGTACCGGT	218		
Db	1	GCTCATCCTGGTCATGGTTACCAAGAAACTGAGAGCATGACGGACAAGTACAGGT	60		
QY	219	GCACCTGTCAAGTGGGTGACCTCCTCTTTGTGTATCATCACTCCCTCTCTGGGCAGTTGATGC	278		
Db	61	GCACCTGTCAAGTGGCCGACCTCCTTTGTTCATCAAGTTCCTCTCTGGGCAGTTGATGC	120		
QY	279	CATGGCTGACTGGTACTTGGGAAATTTTGTGTAAGGCTGCCATATCATCTACACTGT	338		

```
Db 121 CGTGGCAAACTGGTACTTTGGAACTTCCTATGCAAGCAGTCCATGTCATCTACACAGT 180
Qy 339 CAACCTCTACAGCAGCGTTCTATCCTGGCCCTTCATCAGCCTGACCGGTACCTCGCCAT 398
Db 181 CAACCTCTACAGCAGTGTCTCTATCCTGGCCCTTCATCAGTCTGGACCGCTACCTGGCCAT 240
Qy 399 TGTCACGCCCAACCAACAGTCAAGGCCCAAGAACTGCTGGCTGAAAGGCAGTCTATGT 458
Db 241 CGTCCAGGCCCAACCAACAGTCAAGGCCCAAGAACTGCTGGCTGAAAGGTGGTCTATGT 300
Qy 459 GGGGCTCTGATCCAGCCCTCCCTCCGACTATACCTGACTTTCATCTTTTCCGACGTCAG 518
Db 301 TGGGCTCTGGATCCCTGGCCCTCTGCTGACTTATCCCGACTTCATCTTTGCCAAG--- 356
Qy 519 CCAGGGGGACATCAGTCAGGGGGATGACAGGTACATCTGTGACCGCCTTTACCCCGATAG 578
Db 357 -----TCAGTGAGGCAGATGACAGATATATCTGTGACCGCTTCTACCCCAATGA 405
Qy 579 CCTGTGATGGTGGTGTTCATATCCAGCATATATAGTGGGTCTCATCTGCGCCGGCAT 638
Db 406 CTCTGTGGTGGTGTGTTCAGTTTCAGACATCATGTTGGCCTTATCTGCTGGTAT 465
Qy 639 CGTCATCTCTCTGTTACTGATCATCATCTTAAGCTGTCACTTCCAAAGGGCCACCA 698
Db 466 TGTCTATCTGTCTGCTATTGCAATTATCATCTCCAGCTGTCACTCCAAAGGGCCACCA 525
Qy 699 GAAGGCAAGGCCCTCAAGCGACAGTCATCCTCATCTAGCTTTCTTTGCTGCTGGCT 758
Db 526 GAAGGCAAGGCCCTCAAGCACAGTCATCTCATCTGCTTTCTTCCCTGTTGGCT 585
Qy 759 GCCATATTATGTGGGGATCAGCATCGACTCCTTCATCTTTTGGGATCATCAAGCAAGG 818
Db 586 GCCTTACTACATTGGGATCAGCATCGACTCCTTCATCTCTCGAATCATCAAGCAAGG 645
Qy 819 ATGTGACTTCGAGAGCAATTGTGCAAGTGTGATCTCCATCAGAGGCCCTCGCCTTCT 878
Db 646 GTGTGAGTTTGAGAACACTGTGCAAGTGTGATTTCCATCAGGAGGCCCTAGCTTTCTT 705
Qy 879 CCAGTGTGCTGAACCCCATCTCTATGCTTCTCTCGGG-GCCAAAGTTCAAAAGCTCTG 937
Db 706 CCAGTGTGCTGAACCCCATCTCTATGCTTTCTTGGGAGCAATTTTAAACCTCTG 765
Qy 938 CCCAGCATGCACTCACTCCATGAGCAGAGG 968
Db 766 CCCAGCATGCACTCACCTCTGTGAGCAGAGG 796
```

Search completed: May 17, 2004, 16:09:12
Job time : 4894.52 secs